

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-097298

(43)Date of publication of application : 14.04.1998

(51)Int.Cl. G10L 9/18  
G10L 3/00  
G10L 9/14  
H03M 7/30

(21)Application number : 08-251614

(71)Applicant : SONY CORP

(22)Date of filing : 24.09.1996

(72)Inventor : NISHIGUCHI MASAYUKI

IJIMA KAZUYUKI

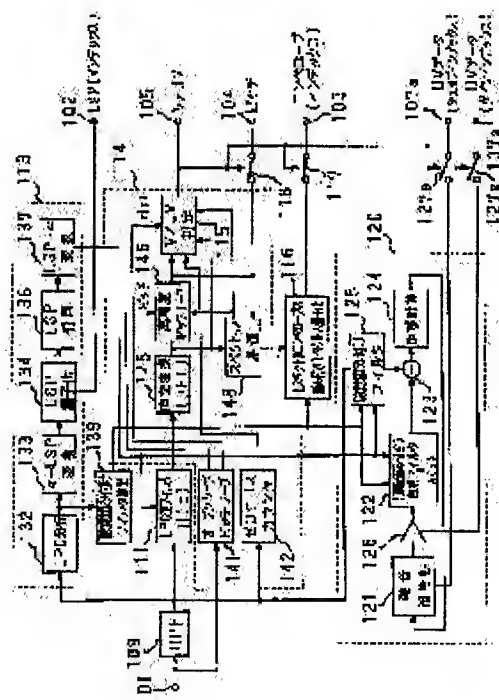
MATSUMOTO ATSUSHI

## (54) VECTOR QUANTIZING METHOD, METHOD AND DEVICE FOR VOICE CODING

### (57)Abstract:

PROBLEM TO BE SOLVED: To reduce the amount of calculation required for the code book search of vector quantization.

SOLUTION: In the case that the data representing the envelope of the harmonic spectrum coming from the spectrum estimating section 148 of a sine wave analyzing/coding section 114 is sent to the vector quantizing section 116 to be vectorially quantized, the similarities between the input vector and all code vectors stored in the code book are determined by approximate calculation to preliminarily select plural code vectors high in similarity. After candidates are focussed to some extent by the preliminary selection simplified in calculation, high-precision regular selection is performed on these candidates alone by regularly selecting the optimum code vector which minimizes the error in reference with the input vector out of the plural code vectors selected by this preliminary selection.



## LEGAL STATUS

[Date of request for examination]

03.03.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3707153

[Date of registration] 12.08.2005

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-97298

(43) 公開日 平成10年(1998) 4月14日

(51) Int.Cl. <sup>6</sup>	識別記号	F I
G 1 0 L 9/18		G 1 0 L 9/18 E
3/00	5 1 5	3/00 5 1 5 B
9/14		9/14 J
		G
H 0 3 M 7/30		H 0 3 M 7/30 B

審査請求 未請求 請求項の数13 O L (全 44 頁)

(21) 出願番号 特願平8-251614

(22) 出願日 平成8年(1996) 9月24日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 西口 正之

東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72) 発明者 飯島 和幸

東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72) 発明者 松本 淳

東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

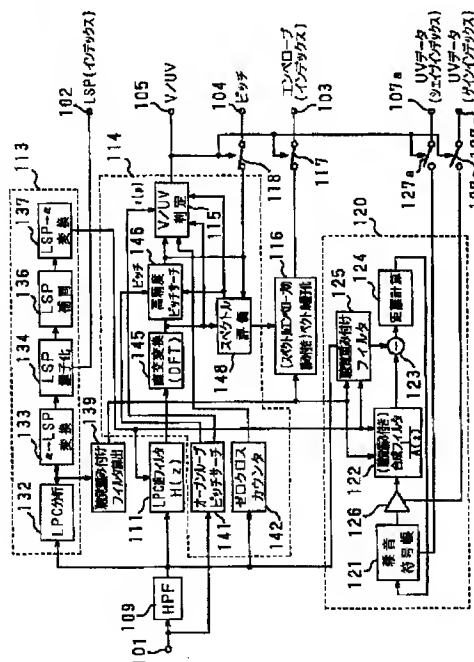
(74) 代理人 弁理士 小池 晃 (外2名)

(54) 【発明の名称】 ベクトル量子化方法、音声符号化方法及び装置

(57) 【要約】

【課題】 ベクトル量子化のコードブックサーチに要する演算量を低減する。

【解決手段】 サイン波分析符号化部114のスペクトル評価部148からのハーモニクススペクトルのエンベロープを表すデータをベクトル量子化部116に送ってベクトル量子化する際に、入力ベクトルと符号帳に蓄えられた全てのコードベクトルとの類似度を近似計算により求めて類似度の高い複数のコードベクトルを予備選択し、この予備選択により選択された複数のコードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを本選択することにより、簡略化された演算の予備選択である程度候補を絞り込んだ後に、これらの候補についてのみ高精度の本選択を行わせる。



## 【特許請求の範囲】

【請求項1】 入力ベクトルを符号帳に蓄えられたコードベクトルと比較して最適のコードベクトルのインデックスを出力するベクトル量子化方法において、上記入力ベクトルと上記符号帳に蓄えられた全てのコードベクトルとの類似度を近似計算により求めて類似度の高い複数のコードベクトルを選択する予備選択工程と、この予備選択工程により選択された複数のコードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを選択する本選択工程とを有することを特徴とするベクトル量子化方法。

【請求項2】 上記符号帳は、複数の符号帳を組み合わせて構成され、最適の組み合わせを構成するコードベクトルが各符号帳毎にそれぞれ選択されることを特徴とする請求項1記載のベクトル量子化方法。

【請求項3】 上記類似度としては、上記入力ベクトルと上記コードベクトルとの内積、重み付き内積、又はこれらを各コードベクトルのノルム又は重み付きノルムで割ったもののいずれかを用いることを特徴とする請求項1記載のベクトル量子化方法。

【請求項4】 上記入力ベクトルは、音声信号に基づく周波数軸上のパラメータのベクトルであり、上記類似度としては、上記周波数軸上のパラメータの低域側にエネルギーが集中し高域側にエネルギーが低下してゆく重みを付けたコードベクトルのノルムで上記コードベクトルの重み付き内積を割ったものを用いることを特徴とする請求項1記載のベクトル量子化方法。

【請求項5】 上記類似度としては、上記周波数軸上のパラメータの低域側にエネルギーが集中し高域側にエネルギーが低下してゆく固定の重みを付けたコードベクトルのノルムで上記コードベクトルの可変の重みを付けた内積を割ったものを用いることを特徴とする請求項4記載のベクトル量子化方法。

【請求項6】 入力音声信号を時間軸上で所定の符号化単位で区分して各符号化単位で符号化を行う音声符号化方法において、

入力音声信号に基づく信号をサイン波分析してハーモニクススペクトルを求める工程と、

上記符号化単位毎の上記ハーモニクススペクトルに基づくパラメータを入力ベクトルとしてベクトル量子化することにより符号化する工程とを有し、

上記ベクトル量子化は、上記入力ベクトルと符号帳に蓄えられた全てのコードベクトルとの類似度を近似計算により求めて類似度の高い複数のコードベクトルを選択する予備選択工程と、この予備選択工程により選択された複数のコードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを選択する本選択工程とを有して成ることを特徴とする音声符号化方法。

【請求項7】 上記符号帳は、複数の符号帳を組み合わ

せて構成され、最適の組み合わせを構成するコードベクトルが各符号帳毎にそれぞれ選択されることを特徴とする請求項6記載の音声符号化方法。

【請求項8】 上記類似度としては、上記入力ベクトルと上記コードベクトルとの内積、重み付き内積、又はこれらを各コードベクトルのノルム又は重み付きノルムで割ったもののいずれかを用いることを特徴とする請求項6記載の音声符号化方法。

【請求項9】 上記ノルムの重み付けに、低域側にエネルギーが集中し高域側にエネルギーが低下してゆく重みを付けたコードベクトルのノルムで上記コードベクトルの重み付き内積を割ったものを用いることを特徴とする請求項8記載の音声符号化方法。

【請求項10】 入力音声信号を時間軸上で所定の符号化単位で区分して各符号化単位で符号化を行う音声符号化装置において、

入力音声信号の短期予測残差を求める予測符号化手段と、

求められた短期予測残差に対してサイン波分析符号化を施すサイン波分析符号化手段とを有し、

上記サイン波分析符号化手段は、サイン波分析により得られたハーモニクススペクトルに基づくパラメータを入力ベクトルとして、この入力ベクトルと符号帳に蓄えられた全てのコードベクトルとの類似度を近似計算により求めて類似度の高い複数のコードベクトルを予備選択し、これらの予備選択された複数のコードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを選択するベクトル量子化手段を有することを特徴とする音声符号化装置。

【請求項11】 上記符号帳は、複数の符号帳を組み合わせて構成され、最適の組み合わせを構成するコードベクトルが各符号帳毎にそれぞれ選択されることを特徴とする請求項10記載の音声符号化装置。

【請求項12】 上記類似度としては、上記入力ベクトルと上記コードベクトルとの内積、重み付き内積、又はこれらを各コードベクトルのノルム又は重み付きノルムで割ったもののいずれかを用いることを特徴とする請求項10記載の音声符号化装置。

【請求項13】 上記ノルムの重み付けに、有声音部分の低域側にエネルギーが集中し高域側にエネルギーが低下してゆく重みを付けたコードベクトルのノルムで上記コードベクトルの重み付き内積を割ったものを用いることを特徴とする請求項12記載の音声符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、入力ベクトルを符号帳に蓄えられたコードベクトルと比較して最適のコードベクトルのインデックスを出力するベクトル量子化方法、入力音声信号をブロックやフレーム等の所定の符号化単位で区分して、区分された符号化単位毎にベクトル

量子化を含んだ符号化処理を行うような音声符号化方法及び装置に関する。

【0002】

【従来の技術】オーディオ信号やビデオ信号等をデジタル化し、圧縮符号化を施す場合に、入力データの複数個をまとめてベクトルとし1つの符号（インデクス）で表現するようなベクトル量子化が知られている。

【0003】このベクトル量子化においては、入力される種々のベクトルの代表的なパターンを予め学習等によって決定して、それぞれに符号（インデクス）を与えて符号帳（コードブック）に蓄えておき、入力ベクトルと符号帳の各パターン（コードベクトル）との比較、すなわちパターンマッチングを行い、最も類似度や相関性の高いパターンの符号を出力する。この類似度や相関性は、入力ベクトルと各コードベクトルとの間の歪尺度や誤差エネルギー等を計算することにより求められ、歪や誤差が小さいほど類似度や相関性が高いことになる。

【0004】ところで、オーディオ信号（音声信号や音響信号を含む）の時間領域や周波数領域における統計的性質と人間の聴感上の特性を利用して信号圧縮を行うような符号化方法が種々知られている。この符号化方法としては、大別して時間領域での符号化、周波数領域での符号化、分析合成符号化等が挙げられる。

【0005】音声信号等の高能率符号化の例として、ハーモニック（Harmonic）符号化、MBE（Multiband Excitation: マルチバンド励起）符号化等のサイン波分析符号化や、SBC（Sub-band Coding: 帯域分割符号化）、LPC（Linear Predictive Coding: 線形予測符号化）、あるいはDCT（離散コサイン変換）、MDC T（モデファイドDCT）、FFT（高速フーリエ変換）等が知られている。

【0006】このような音声信号等の高能率符号化において、例えば得られたハーモニクススペクトル等のパラメータに対して、上述したようなベクトル量子化が採用されている。

【0007】

【発明が解決しようとする課題】ところで、符号帳に蓄えられているパターン、すなわちコードベクトルの個数が多い場合や、複数の符号帳を組み合わせで構成される多段のベクトル量子化器の場合には、上記パターンマッチングの際のコードベクトルの検索（サーチ）の回数が多くなり、演算量が増大するという欠点がある。特に、複数の符号帳を組み合わせる場合には、各符号帳のコードベクトルの個数の積の回数の類似度演算が必要となるため、コードブックサーチの演算量がかなり大きくなってしまふ。

【0008】本発明は、このような実情に鑑みてなされたものであり、コードブックサーチの演算量を抑えることが可能なベクトル量子化方法、音声符号化方法及び装置の提供を目的とする。

【0009】

【課題を解決するための手段】本発明に係るベクトル量子化方法は、上述した課題を解決するために、ベクトル量子化しようとする入力ベクトルと符号帳（コードブック）に蓄えられた全てのコードベクトルとの類似度を近似計算により求めて類似度の高い複数のコードベクトルを予備選択し、この予備選択により選択された複数のコードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを本選択することを特徴とする。

【0010】この予備選択を行った後に本選択を行うことにより、簡略化された演算の予備選択である程度候補を絞り込んだ後に、これらの候補についてのみ高精度の本選択を行わせることができ、符号帳検索（コードブックサーチ）に要する演算量を低減できる。

【0011】ここで上記符号帳は、複数の符号帳を組み合わせで構成され、最適の組み合わせを構成するコードベクトルが各符号帳毎にそれぞれ選択されることが挙げられる。また、上記類似度としては、上記入力ベクトルと上記コードベクトルとの内積、重み付き内積、又はこれらを各コードベクトルのノルム又は重み付きノルムで割ったもののいずれかを用いることが挙げられる。

【0012】また、本発明に係る音声符号化方法は、入力音声信号又は入力音声信号の短期予測残差をサイン波分析してハーモニクススペクトルを求め、符号化単位毎の上記ハーモニクススペクトルに基づくパラメータを入力ベクトルとしてベクトル量子化することにより符号化する音声符号化方法であって、上記ベクトル量子化は、上記入力ベクトルと符号帳に蓄えられた全てのコードベクトルとの類似度を近似計算により求めて類似度の高い複数のコードベクトルを予備選択し、この予備選択により選択された複数のコードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを本選択することを特徴とする。

【0013】上記類似度としては、上記入力ベクトルと上記コードベクトルとの内積、重み付き内積、又はこれらを各コードベクトルのノルム又は重み付きノルムで割ったもののいずれかを用い、また上記ノルムの重み付けには、低域側にエネルギーが集中し高域側にエネルギーが低下してゆく重みを用い、そのような重みを付けたコードベクトルのノルムで上記コードベクトルの重み付き内積を割ったものを用いて、上記類似度を求めることが挙げられる。

【0014】このような音声符号化方法が適用された音声符号化装置を構成できる。

【0015】

【発明の実施の形態】以下、本発明に係る好ましい実施の形態について説明する。先ず、図1は、本発明に係るベクトル量子化方法の実施の形態が適用された音声符号化装置の基本構成を示している。

【0016】ここで、図1の音声信号符号化装置の基本的な考え方は、入力音声信号の短期予測残差例えばLPC（線形予測符号化）残差を求めてサイン波分析（sinusoidal analysis）符号化、例えばハーモニックコーディング（harmonic coding）を行う第1の符号化部110と、入力音声信号に対して位相再現性のある波形符号化により符号化する第2の符号化部120とを有し、入力信号の有声音（V：Voiced）の部分の符号化に第1の符号化部110を用い、入力信号の無声音（UV：Unvoiced）の部分の符号化には第2の符号化部120を用い

10 るようにすることである。  
【0017】上記第1の符号化部110には、例えばLPC残差をハーモニック符号化やマルチバンド励起（MBE）符号化のようなサイン波分析符号化を行う構成が用いられる。上記第2の符号化部120には、例えば合成による分析法を用いて最適ベクトルのクローズドループサーチによるベクトル量子化を用いた符号励起線形予測（CELP）符号化の構成が用いられる。

【0018】図1の例では、入力端子101に供給された音声信号が、第1の符号化部110のLPC逆フィルタ111及びLPC分析・量子化部113に送られている。LPC分析・量子化部113から得られたLPC係数あるいはいわゆる $\alpha$ パラメータは、LPC逆フィルタ111に送られて、このLPC逆フィルタ111により入力音声信号の線形予測残差（LPC残差）が取り出される。また、LPC分析・量子化部113からは、後述するようにLSP（線スペクトル対）の量子化出力が取り出され、これが出力端子102に送られる。LPC逆フィルタ111からのLPC残差は、サイン波分析符号化部114に送られる。サイン波分析符号化部114では、ピッチ検出やスペクトルエンベロープ振幅計算が行われると共に、V（有声音）/UV（無声音）判定部115によりV/UVの判定が行われる。サイン波分析符号化部114からのスペクトルエンベロープ振幅データがベクトル量子化部116に送られる。スペクトルエンベロープのベクトル量子化出力としてのベクトル量子化部116からのコードブックインデックスは、スイッチ117を介して出力端子103に送られ、サイン波分析符号化部114からの出力は、スイッチ118を介して出力端子104に送られる。また、V/UV判定部115からのV/UV判定出力は、出力端子105に送られると共に、スイッチ117、118の制御信号として送られており、上述した有声音（V）のとき上記インデックス及びピッチが選択されて各出力端子103及び104からそれぞれ取り出される。

【0019】図1の第2の符号化部120は、この例ではCELP（符号励起線形予測）符号化構成を有しており、雑音符号帳121からの出力を、重み付きの合成フィルタ122により合成処理し、得られた重み付き音声

声信号を聴覚重み付けフィルタ125を介して得られた音声との誤差を取り出し、この誤差を距離計算回路124に送って距離計算を行い、誤差が最小となるようなベクトルを雑音符号帳121でサーチするような、合成による分析（Analysis by Synthesis）法を用いたクローズドループサーチを用いた時間軸波形のベクトル量子化を行っている。このCELP符号化は、上述したように無声音部分の符号化に用いられており、雑音符号帳121からのUVデータとしてのコードブックインデックスは、上記V/UV判定部115からのV/UV判定結果が無声音（UV）のときオンとなるスイッチ127を介して、出力端子107より取り出される。

【0020】次に、図2は、本発明に係る音声復号化方法の一実施の形態が適用された音声信号復号化装置として、上記図1の音声信号符号化装置に対応する音声信号復号化装置の基本構成を示すブロック図である。

【0021】この図2において、入力端子202には上記図1の出力端子102からの上記LSP（線スペクトル対）の量子化出力としてのコードブックインデックスが10 入力される。入力端子203、204、及び205には、上記図1の各出力端子103、104、及び105からの各出力、すなわちエンベロープ量子化出力としてのインデックス、ピッチ、及びV/UV判定出力がそれぞれ入力される。また、入力端子207には、上記図1の出力端子107からのUV（無声音）用のデータとしてのインデックスが入力される。

【0022】入力端子203からのエンベロープ量子化出力としてのインデックスは、逆ベクトル量子化器212に送られて逆ベクトル量子化され、LPC残差のスペクトルエンベロープが求められて有声音合成部211に送られる。有声音合成部211は、サイン波合成により有声音部分のLPC（線形予測符号化）残差を合成するものであり、この有声音合成部211には入力端子204及び205からのピッチ及びV/UV判定出力も供給されている。有声音合成部211からの有声音のLPC残差は、LPC合成フィルタ214に送られる。また、入力端子207からのUVデータのインデックスは、無声音合成部220に送られて、雑音符号帳を参照することにより無声音部分のLPC残差が取り出される。このLPC残差もLPC合成フィルタ214に送られる。LPC合成フィルタ214では、上記有声音部分のLPC残差と無声音部分のLPC残差とがそれぞれ独立に、LPC合成処理が施される。あるいは、有声音部分のLPC残差と無声音部分のLPC残差とが加算されたものに対してLPC合成処理を施すようにしてもよい。ここで入力端子202からのLSPのインデックスは、LPCパラメータ再生部213に送られて、LPCの $\alpha$ パラメータが取り出され、これがLPC合成フィルタ214に送られる。LPC合成フィルタ214によりLPC合成されて得られた音声信号は、出力端子201より取り出され

る。

【0023】次に、上記図1に示した音声信号符号化装置のより具体的な構成について、図3を参照しながら説明する。なお、図3において、上記図1の各部と対応する部分には同じ指示符号を付している。

【0024】この図3に示された音声信号符号化装置において、入力端子101に供給された音声信号は、ハイパスフィルタ(HPF)109にて不要な帯域の信号を除去するフィルタ処理が施された後、LPC(線形予測符号化)分析・量子化部113のLPC分析回路132と、LPC逆フィルタ回路111とに送られる。

【0025】LPC分析・量子化部113のLPC分析回路132は、入力信号波形の256サンプル程度の長さを符号化単位の1ブロックとしてハミング窓をかけて、自己相関法により線形予測係数、いわゆる $\alpha$ パラメータを求める。データ出力の単位となるフレーミングの間隔は、160サンプル程度とする。サンプリング周波数 $f_s$ が例えば8kHzのとき、1フレーム間隔は160サンプルで20msecとなる。

【0026】LPC分析回路132からの $\alpha$ パラメータは、 $\alpha \rightarrow$ LSP変換回路133に送られて、線スペクトル対(LSP)パラメータに変換される。これは、直接型のフィルタ係数として求めた $\alpha$ パラメータを、例えば10個、すなわち5対のLSPパラメータに変換する。変換は例えばニュートン-ラプソン法等を用いて行う。このLSPパラメータに変換するのは、 $\alpha$ パラメータよりも補間特性に優れているからである。

【0027】 $\alpha \rightarrow$ LSP変換回路133からのLSPパラメータは、LSP量子化器134によりマトリクスあるいはベクトル量子化される。このとき、フレーム間差をとってからベクトル量子化してもよく、複数フレーム分をまとめてマトリクス量子化してもよい。ここでは、20msecを1フレームとし、20msec毎に算出されるLSPパラメータを2フレーム分まとめて、マトリクス量子化及びベクトル量子化している。

【0028】このLSP量子化器134からの量子化出力、すなわちLSP量子化のインデックスは、端子102を介して取り出され、また量子化済みのLSPベクトルは、LSP補間回路136に送られる。

【0029】LSP補間回路136は、上記20msecあるいは40msec毎に量子化されたLSPのベクトルを補間し、8倍のレートにする。すなわち、2.5msec毎にLSPベクトルが更新されるようにする。これは、残差波形をハーモニック符号化復号化方法により分析合成すると、その合成波形のエンベロープは非常になだらかでスムーズな波形になるため、LPC係数が20msec毎に急激に変化すると異音を発生することがあるからである。すなわち、2.5msec毎にLPC係数が徐々に変化してゆくようにすれば、このような異音の発生を防ぐことができる。

【0030】このような補間が行われた2.5msec毎のLSPベクトルを用いて入力音声の逆フィルタリングを実行するために、LSP $\rightarrow\alpha$ 変換回路137により、LSPパラメータを例えば10次程度の直接型フィルタの係数である $\alpha$ パラメータに変換する。このLSP $\rightarrow\alpha$ 変換回路137からの出力は、上記LPC逆フィルタ回路111に送られ、このLPC逆フィルタ111では、2.5msec毎に更新される $\alpha$ パラメータにより逆フィルタリング処理を行って、滑らかな出力を得るようにしている。このLPC逆フィルタ111からの出力は、サイン波分析符号化部114、具体的には例えばハーモニック符号化回路、の直交変換回路145、例えばDFT(離散フーリエ変換)回路に送られる。

【0031】LPC分析・量子化部113のLPC分析回路132からの $\alpha$ パラメータは、聴覚重み付けフィルタ算出回路139に送られて聴覚重み付けのためのデータが求められ、この重み付けデータが後述する聴覚重み付きのベクトル量子化器116と、第2の符号化部120の聴覚重み付けフィルタ125及び聴覚重み付きの合成フィルタ122とに送られる。

【0032】ハーモニック符号化回路等のサイン波分析符号化部114では、LPC逆フィルタ111からの出力を、ハーモニック符号化の方法で分析する。すなわち、ピッチ検出、各ハーモニクスの振幅 $A_m$ の算出、有声音(V)/無声音(UV)の判別を行い、ピッチによって変化するハーモニクスのエンベロープあるいは振幅 $A_m$ の個数を次元変換して一定数にしている。

【0033】図3に示すサイン波分析符号化部114の具体例においては、一般のハーモニック符号化を想定しているが、特に、MBE(Multiband Excitation: マルチバンド励起)符号化の場合には、同時刻(同じブロックあるいはフレーム内)の周波数軸領域いわゆるバンド毎に有声音(Voiced)部分と無声音(Unvoiced)部分とが存在するという仮定でモデル化することになる。それ以外のハーモニック符号化では、1ブロックあるいはフレーム内の音声の有声音か無声音かの択一的な判定がなされることになる。なお、以下の説明中のフレーム毎のV/UVとは、MBE符号化に適用した場合には全バンドがUVのときを当該フレームのUVとしている。ここで上記MBEの分析合成手法については、本件出願人が先に提案した特願平4-91422号明細書及び図面に詳細な具体例を開示している。

【0034】図3のサイン波分析符号化部114のオープンループピッチサーチ部141には、上記入力端子101からの入力音声信号が、またゼロクロスカウンタ142には、上記HPF(ハイパスフィルタ)109からの信号がそれぞれ供給されている。サイン波分析符号化部114の直交変換回路145には、LPC逆フィルタ111からのLPC残差あるいは線形予測残差が供給されている。オープンループピッチサーチ部141では、

入力信号のLPC残差をとってオープンループによる比較的ラフなピッチのサーチが行われ、抽出された粗ピッチデータは高精度ピッチサーチ146に送られて、後述するようなクロズドループによる高精度のピッチサーチ（ピッチのファインサーチ）が行われる。また、オープンループピッチサーチ部141からは、上記粗ピッチデータと共にLPC残差の自己相関の最大値をパワーで正規化した正規化自己相関最大値 $r(p)$ が取り出され、V/UV（有声音/無声音）判定部115に送られている。

【0035】直交変換回路145では例えばDFT（離散フーリエ変換）等の直交変換処理が施されて、時間軸上のLPC残差が周波数軸上のスペクトル振幅データに変換される。この直交変換回路145からの出力は、高精度ピッチサーチ部146及びスペクトル振幅あるいはエンベロープを評価するためのスペクトル評価部148に送られる。

【0036】高精度（ファイン）ピッチサーチ部146には、オープンループピッチサーチ部141で抽出された比較的ラフな粗ピッチデータと、直交変換部145により例えばDFTされた周波数軸上のデータとが供給されている。この高精度ピッチサーチ部146では、上記粗ピッチデータ値を中心に、0.2～0.5きざみで±数サンプルずつ振って、最適な小数点付き（フローティング）のファインピッチデータの値へ追い込む。このときのファインサーチの手法として、いわゆる合成による分析（Analysis by Synthesis）法を用い、合成されたパワースペクトルが原音のパワースペクトルに最も近くなるようにピッチを選んでいる。このようなクロズドループによる高精度のピッチサーチ部146からのピッチデータについては、スイッチ118を介して出力端子104に送っている。

【0037】スペクトル評価部148では、LPC残差の直交変換出力としてのスペクトル振幅及びピッチに基づいて各ハーモニクスの大きさ及びその集合であるスペクトルエンベロープが評価され、高精度ピッチサーチ部146、V/UV（有声音/無声音）判定部115及び聴覚重み付きのベクトル量子化器116に送られる。

【0038】V/UV（有声音/無声音）判定部115は、直交変換回路145からの出力と、高精度ピッチサーチ部146からの最適ピッチと、スペクトル評価部148からのスペクトル振幅データと、オープンループピッチサーチ部141からの正規化自己相関最大値 $r(p)$ と、ゼロクロスカウンタ142からのゼロクロスカウンタ値とに基づいて、当該フレームのV/UV判定が行われる。さらに、MBEの場合の各バンド毎のV/UV判定結果の境界位置も当該フレームのV/UV判定の一条件としてもよい。このV/UV判定部115からの判定出力は、出力端子105を介して取り出される。

【0039】ところで、スペクトル評価部148の出力

部あるいはベクトル量子化器116の入力部には、データ数変換（一種のサンプリングレート変換）部が設けられている。このデータ数変換部は、上記ピッチに応じて周波数軸上での分割帯域数が異なり、データ数が異なることを考慮して、エンベロープの振幅データ $|A_n|$ を一定の個数にするためのものである。すなわち、例えば有効帯域を3400kHzまでとすると、この有効帯域が上記ピッチに応じて、8バンド～63バンドに分割されることになり、これらの各バンド毎に得られる上記振幅データ $|A_n|$ の個数 $m_{nx}+1$ も8～63と変化することになる。このためデータ数変換部119では、この可変個数 $m_{nx}+1$ の振幅データを一定個数M個、例えば44個、のデータに変換している。

【0040】このスペクトル評価部148の出力部あるいはベクトル量子化器116の入力部に設けられたデータ数変換部からの上記一定個数M個（例えば44個）の振幅データあるいはエンベロープデータが、ベクトル量子化器116により、所定個数、例えば44個のデータ毎にまとめられてベクトルとされ、重み付きベクトル量子化が施される。この重みは、聴覚重み付けフィルタ算出回路139からの出力により与えられる。ベクトル量子化器116からの上記エンベロープのインデックスは、スイッチ117を介して出力端子103より取り出される。なお、上記重み付きベクトル量子化に先だって、所定個数のデータから成るベクトルについて適当なリーク係数を用いたフレーム間差分をとっておくようにしてもよい。

【0041】次に、第2の符号化部120について説明する。第2の符号化部120は、いわゆるCELP（符号励起線形予測）符号化構成を有しており、特に、入力音声信号の無声音部分の符号化のために用いられている。この無声音部分用のCELP符号化構成において、雑音符号帳、いわゆるストキャスティック・コードブック（stochastic code book）121からの代表値出力である無声音のLPC残差に相当するノイズ出力を、ゲイン回路126を介して、聴覚重み付きの合成フィルタ122に送っている。重み付きの合成フィルタ122では、入力されたノイズをLPC合成処理し、得られた重み付き無声音の信号を減算器123に送っている。減算器123には、上記入力端子101からHPF（ハイパスフィルタ）109を介して供給された音声信号を聴覚重み付けフィルタ125で聴覚重み付けした信号が入力されており、合成フィルタ122からの信号との差分あるいは誤差を取り出している。なお、聴覚重み付けフィルタ125の出力から聴覚重み付き合成フィルタの零入力応答を事前に差し引いておくものとする。この誤差を距離計算回路124に送って距離計算を行い、誤差が最小となるような代表値ベクトルを雑音符号帳121でサーチする。このような合成による分析（Analysis by Synthesis）法を用いたクロズドループサーチを用いた

10

20

30

40

50



時間軸波形のベクトル量子化を行っている。

【0042】このCELP符号化構成を用いた第2の符号化部120からのUV（無声音）部分用のデータとしては、雑音符号帳121からのコードブックのシェイプインデックスと、ゲイン回路126からのコードブックのゲインインデックスとが取り出される。雑音符号帳121からのUVデータであるシェイプインデックスは、スイッチ127sを介して出力端子107sに送られ、ゲイン回路126のUVデータであるゲインインデックスは、スイッチ127gを介して出力端子107gに送られている。

【0043】ここで、これらのスイッチ127s、127g及び上記スイッチ117、118は、上記V/UV判定部115からのV/UV判定結果によりオン/オフ制御され、スイッチ117、118は、現在伝送しようとするフレームの音声信号のV/UV判定結果が有声音（V）のときオンとなり、スイッチ127s、127gは、現在伝送しようとするフレームの音声信号が無声音（UV）のときオンとなる。

【0044】次に、図4は、上記図2に示した本発明に係る実施の形態としての音声信号復号化装置のより具体的な構成を示している。この図4において、上記図2の各部と対応する部分には、同じ指示符号を付している。

【0045】この図4において、入力端子202には、上記図1、3の出力端子102からの出力に相当するLSPのベクトル量子化出力、いわゆるコードブックのインデックスが供給されている。

【0046】このLSPのインデックスは、LPCパラメータ再生部213のLSPの逆ベクトル量子化器231に送られてLSP（線スペクトル対）データに逆ベクトル量子化され、LSP補間回路232、233に送られてLSPの補間処理が施された後、LSP→ $\alpha$ 変換回路234、235でLPC（線形予測符号）の $\alpha$ パラメータに変換され、この $\alpha$ パラメータがLPC合成フィルタ214に送られる。ここで、LSP補間回路232及びLSP→ $\alpha$ 変換回路234は有声音（V）用であり、LSP補間回路233及びLSP→ $\alpha$ 変換回路235は無声音（UV）用である。またLPC合成フィルタ214は、有声音部分のLPC合成フィルタ236と、無声音部分のLPC合成フィルタ237とを分離している。すなわち、有声音部分と無声音部分とでLPCの係数補間を独立に行うようにして、有声音から無声音への遷移部や、無声音から有声音への遷移部で、全く性質の異なるLSP同士を補間することによる悪影響を防止している。

【0047】また、図4の入力端子203には、上記図1、図3のエンコーダ側の端子103からの出力に対応するスペクトルエンベロープ（Am）の重み付けベクトル量子化されたコードインデックスデータが供給され、入力端子204には、上記図1、図3の端子104からの

ピッチのデータが供給され、入力端子205には、上記図1、図3の端子105からのV/UV判定データが供給されている。

【0048】入力端子203からのスペクトルエンベロープAmのベクトル量子化されたインデックスデータは、逆ベクトル量子化器212に送られて逆ベクトル量子化が施され、上記データ数変換に対応する逆変換が施されて、スペクトルエンベロープのデータとなって、有声音合成部211のサイン波合成回路215に送られている。

【0049】なお、エンコード時にスペクトルのベクトル量子化に先だってフレーム間差分をとっている場合には、ここでの逆ベクトル量子化後にフレーム間差分の復号を行ってからデータ数変換を行い、スペクトルエンベロープのデータを得る。

【0050】サイン波合成回路215には、入力端子204からのピッチ及び入力端子205からの上記V/UV判定データが供給されている。サイン波合成回路215からは、上述した図1、図3のLPC逆フィルタ111からの出力に相当するLPC残差データが取り出され、これが加算器218に送られている。このサイン波合成の具体的な手法については、例えば本件出願人が先に提案した、特願平4-91422号の明細書及び図面、あるいは特願平6-198451号の明細書及び図面に開示されている。

【0051】また、逆ベクトル量子化器212からのエンベロープのデータと、入力端子204、205からのピッチ、V/UV判定データとは、有声音（V）部分のノイズ加算のためのノイズ合成回路216に送られている。このノイズ合成回路216からの出力は、重み付き重量加算回路217を介して加算器218に送っている。これは、サイン波合成によって有声音のLPC合成フィルタへの入力となるエクサイテイション（Excitation: 励起、励振）を作ると、男声等の低いピッチの音で鼻づまり感がある点、及びV（有声音）とUV（無声音）とで音質が急激に変化し不自然に感じる場合がある点を考慮し、有声音部分のLPC合成フィルタ入力すなわちエクサイテイションについて、音声符号化データに基づくパラメータ、例えばピッチ、スペクトルエンベロープ振幅、フレーム内の最大振幅、残差信号のレベル等を考慮したノイズをLPC残差信号の有声音部分に加えているものである。

【0052】加算器218からの加算出力は、LPC合成フィルタ214の有声音用の合成フィルタ236に送られてLPCの合成処理が施されることにより時間波形データとなり、さらに有声音用ポストフィルタ238vでフィルタ処理された後、加算器239に送られる。

【0053】次に、図4の入力端子207s及び207gには、上記図3の出力端子107s及び107gからのUVデータとしてのシェイプインデックス及びゲインイ

ンデックスがそれぞれ供給され、無声音合成部220に送られている。端子207sからのシェイプインデックスは、無声音合成部220の雑音符号帳221に、端子207gからのゲインインデックスはゲイン回路222にそれぞれ送られている。雑音符号帳221から読み出された代表値出力は、無声音のLPC残差に相当するノイズ信号成分であり、これがゲイン回路222で所定のゲインの振幅となり、窓かけ回路223に送られて、上記有声音部分とのつながりを円滑化するための窓かけ処理が施される。

【0054】窓かけ回路223からの出力は、無声音合成部220からの出力として、LPC合成フィルタ214のUV（無声音）用の合成フィルタ237に送られる。合成フィルタ237では、LPC合成処理が施されることにより無声音部分の時間波形データとなり、この無声音部分の時間波形データは無声音用ポストフィルタ238uでフィルタ処理された後、加算器239に送ら\*

＊れる。

【0055】加算器239では、有声音用ポストフィルタ238vからの有声音部分の時間波形信号と、無声音用ポストフィルタ238uからの無声音部分の時間波形データとが加算され、出力端子201より取り出される。

【0056】ところで、上記音声信号符号化装置では、要求される品質に合わせ異なるビットレートの出力データを出力することができ、出力データのビットレートが可変されて出力される。

【0057】具体的には、出力データのビットレートを、低ビットレートと高ビットレートとに切り換えることができる。例えば、低ビットレートを2 kbpsとし、高ビットレートを6 kbpsとする場合には、以下の表1に示す各ビットレートのデータが出力される。

【0058】

【表1】

	2Kbps		6Kbps	
V/UV判定出力	1bit/20msec		1bit/20msec	
LSP量子化インデックス	32bits/40msec		48bits/40msec	
有声音時 (V)	ピッチデータ	8bits/20msec	ピッチデータ	8bits/20msec
	インデックス	8bits/20msec	インデックス	8bits/20msec
	シェイプ(1段目)	5+5bits/20msec	シェイプ(1段目)	5+5bits/20msec
	ゲイン	5bits/20msec	ゲイン	5bits/20msec
無声音時 (UV)	インデックス	11bits/10msec	インデックス	23bits/5msec
	シェイプ(1段目)	7bits/10msec	シェイプ(1段目)	9bits/5msec
	ゲイン	4bits/10msec	ゲイン	6bits/5msec
	シェイプ(2段目)		シェイプ(2段目)	5bits/5msec
有声音時 無声音時	40bits/20msec		120bits/20msec	
	39bits/20msec		117bits/20msec	

【0059】出力端子104からのピッチデータについては、有声音時に、常に8bits/20msecで出力され、出力端子105から出力されるV/UV判定出力は、常に1bit/20msecである。出力端子102から出力されるLSP量子化のインデックスは、32bits/40msecと48bits/40msecとの間で切り換えが行われる。また、出力端子103から出力される有声音時(V)のインデックスは、15bits/20msecと87bits/20msecとの間で切り換えが行われ、出力端子107s、107gから出力される無声音時(UV)のインデックスは、11bits/10msecと23bits/5msecとの間で切り換えが行われる。これにより、有声音時(V)の出力データは、2 kbpsでは40bits/20msecとなり、6 kbpsでは120bits/20msecとなる。

また、無声音時(UV)の出力データは、2 kbpsでは39bits/20msecとなり、6 kbpsでは117bits/20msecとなる。

【0060】尚、上記LSP量子化のインデックス、有声音時(V)のインデックス、及び無声音時(UV)のインデックスについては、後述する各部の構成と共に説明する。

【0061】次に、図5及び図6を用いて、LSP量子化器134におけるマトリクス量子化及びベクトル量子化について詳細に説明する。

【0062】上述のように、LPC分析回路132からの $\alpha$ パラメータは、 $\alpha \rightarrow$  LSP変換回路133に送られて、LSPパラメータに変換される。例えば、LPC分析回路132でP次のLPC分析を行う場合には、 $\alpha$ パ

ラメータはP個算出される。このP個の $\alpha$ パラメータは、LSPパラメータに変換され、バッファ610に保持される。

【0063】このバッファ610からは、2フレーム分のLSPパラメータが出力される。2フレーム分のLSPパラメータはマトリクス量子化部620でマトリクス量子化される。マトリクス量子化部620は、第1のマトリクス量子化部620<sub>1</sub>と第2のマトリクス量子化部620<sub>2</sub>とから成る。2フレーム分のLSPパラメータは、第1のマトリクス量子化部620<sub>1</sub>でマトリクス量子化され、これにより得られる量子化誤差が、第2のマトリクス量子化部620<sub>2</sub>でさらにマトリクス量子化される。これらのマトリクス量子化により、時間軸方向及び周波数軸方向の相関を取り除く。

【0064】マトリクス量子化部620<sub>2</sub>からの2フレーム分の量子化誤差は、ベクトル量子化部640に入力される。ベクトル量子化部640は、第1のベクトル量子化部640<sub>1</sub>と第2のベクトル量子化部640<sub>2</sub>とから成る。さらに、第1のベクトル量子化部640<sub>1</sub>は、2つのベクトル量子化部650、660から成り、第2のベクトル量子化部640<sub>2</sub>は、2つのベクトル量子化部670、680から成る。第1のベクトル量子化部640<sub>1</sub>のベクトル量子化部650、660で、マトリクス量子化部620からの量子化誤差が、それぞれ1フレーム毎にベクトル量子化される。これにより得られる量子化誤差ベクトルは、第2のベクトル量子化部640<sub>2</sub>のベクトル量子化部670、680で、さらにベクトル\*

\*量子化される。これらのベクトル量子化により、周波数軸方向の相関を処理する。

【0065】このように、マトリクス量子化を施す工程を行うマトリクス量子化部620は、第1のマトリクス量子化工程を行う第1のマトリクス量子化部620<sub>1</sub>と、この第1のマトリクス量子化による量子化誤差をマトリクス量子化する第2のマトリクス量子化工程を行う第2のマトリクス量子化部620<sub>2</sub>とを少なくとも有し、上記ベクトル量子化を施す工程を行うベクトル量子化部640は、第1のベクトル量子化工程を行う第1のベクトル量子化部640<sub>1</sub>と、この第1のベクトル量子化の際の量子化誤差ベクトルをベクトル量子化する第2のベクトル量子化工程を行う第2のベクトル量子化部640<sub>2</sub>とを少なくとも有する。

【0066】次に、マトリクス量子化及びベクトル量子化について具体的に説明する。

【0067】バッファ610に保持された、2フレーム分のLSPパラメータ、すなわち10×2の行列は、マトリクス量子化器620<sub>1</sub>に送られる。上記第1のマトリクス量子化部620<sub>1</sub>では、2フレーム分のLSPパラメータが加算器621を介して重み付き距離計算器623に送られ、最小となる重み付き距離が算出される。

【0068】この第1のマトリクス量子化部620<sub>1</sub>によるコードブックサーチ時の歪尺度 $d_{wq1}$ は、LSPパラメータ $X_1$ 、量子化値 $X_1'$ を用い、(1)式で示す。

【0069】

【数1】

$$d_{MQ1}(X_1, X_1') = \sum_{n=0}^1 \sum_{i=1}^P w(t, i) (x_1(t, i) - x_1'(t, i))^2 \quad \dots(1)$$

【0070】ここで、tはフレーム番号、iはP次元の番号を示す。

【0071】また、このときの、周波数軸方向及び時間軸方向に重みの制限を考慮しない場合の重みwを(2)式で示す。

【0072】

【数2】

$$w(t, i) = \frac{1}{x(t, i+1) - x(t, i)} + \frac{1}{x(t, i) - x(t, i-1)} \quad \dots(2)$$

ただしtにかかわらず $x(t, 0)=0$ 、 $x(t, p+1)=\pi$

【0073】この(2)式の重みwは、後段のマトリクス量子化及びベクトル量子化でも用いられる。

【0074】算出された重み付き距離はマトリクス量子化器(MQ<sub>1</sub>)622に送られて、マトリクス量子化が行われる。このマトリクス量子化により出力される8ビットのインデックスは信号切換器690に送られる。また、マトリクス量子化による量子化値は、加算器621で、バッファ610からの2フレーム分のLSPパラメータから減算される。重み付き距離計算器623では、

加算器621からの出力を用いて、重み付き距離が算出される。このように、2フレーム毎に、順次、重み付き距離計算器623では重み付き距離が算出されて、マトリクス量子化器622でマトリクス量子化が行われる。重み付き距離が最小となる量子化値が選ばれる。また、加算器621からの出力は、第2のマトリクス量子化部620<sub>2</sub>の加算器631に送られる。

【0075】第2のマトリクス量子化部620<sub>2</sub>でも第1のマトリクス量子化部620<sub>1</sub>と同様にして、マトリクス量子化を行う。上記加算器621からの出力は、加算器631を介して重み付き距離計算器633に送られ、最小となる重み付き距離が算出される。

【0076】この第2のマトリクス量子化部620<sub>2</sub>によるコードブックサーチ時の歪尺度 $d_{wq2}$ を、第1のマトリクス量子化部620<sub>1</sub>からの量子化誤差 $X_2$ 、量子化値 $X_2'$ により、(3)式で示す。

【0077】

【数3】

$$d_{MQ2}(X_2, X'_2) = \sum_{i=0}^1 \sum_{l=1}^P w(l, i) (x_2(l, i) - x'_2(l, i))^2 \quad \dots (3)$$

【0078】この重み付き距離はマトリクス量子化器(MQ<sub>2</sub>) 632に送られて、マトリクス量子化が行われる。このマトリクス量子化により出力される8ビットのインデックスは信号切換器690に送られる。また、マトリクス量子化による量子化値は、加算器631で、2フレーム分の量子化誤差から減算される。重み付き距離計算器633では、加算器631からの出力を用いて、重み付き距離が順次算出されて、重み付き距離が最小となる量子化値が選ばれる。また、加算器631からの出力は、第1のベクトル量子化部640<sub>1</sub>の加算器651、661に1フレームずつ送られる。

\*

$$d_{VQ1}(x_{3-1}, x'_{3-1}) = \sum_{i=1}^P w(0, i) (x_{3-1}(0, i) - x'_{3-1}(0, i))^2 \quad \dots (4)$$

$$d_{VQ2}(x_{3-2}, x'_{3-2}) = \sum_{i=1}^P w(1, i) (x_{3-2}(1, i) - x'_{3-2}(1, i))^2 \quad \dots (5)$$

【0082】この重み付き距離はベクトル量子化器(VQ<sub>1</sub>) 652、ベクトル量子化器(VQ<sub>2</sub>) 662にそれぞれ送られて、ベクトル量子化が行われる。このベクトル量子化により出力される各8ビットのインデックスは信号切換器690に送られる。また、ベクトル量子化による量子化値は、加算器651、661で、入力された2フレーム分の量子化誤差ベクトルから減算される。重み付き距離計算器653、663では、加算器651、661からの出力を用いて、重み付き距離が順次算出されて、重み付き距離が最小となる量子化値が選ばれる。また、加算器651、661からの出力は、第2のベク

トル量子化部640<sub>2</sub>の加算器671、681にそれぞれ送られる。

【0083】ここで、

$$x_{4-1} = x_{3-1} - x'_{3-1}$$

$$x_{4-2} = x_{3-2} - x'_{3-2}$$

と表すときの、この第2のベクトル量子化部640<sub>2</sub>のベクトル量子化器672、682によるコードブックサーチ時の歪尺度 $d_{vq3}$ 、 $d_{vq4}$ を、(6)、(7)式で示す。

【0084】

【数5】

$$d_{VQ3}(x_{4-1}, x'_{4-1}) = \sum_{i=1}^P w(0, i) (x_{4-1}(0, i) - x'_{4-1}(0, i))^2 \quad \dots (6)$$

$$d_{VQ4}(x_{4-2}, x'_{4-2}) = \sum_{i=1}^P w(1, i) (x_{4-2}(1, i) - x'_{4-2}(1, i))^2 \quad \dots (7)$$

【0085】この重み付き距離はベクトル量子化器(VQ<sub>3</sub>) 672、ベクトル量子化器(VQ<sub>4</sub>) 682にそれぞれ送られて、ベクトル量子化が行われる。このベクトル量子化により出力される各8ビットのインデックスは信号切換器690に送られる。また、ベクトル量子化による量子化値は、加算器671、681で、入力された2フレーム分の量子化誤差ベクトルから減算される。重み付き距離計算器673、683では、加算器671、681からの出力を用いて、重み付き距離が順次算出されて、重み付き距離が最小となる量子化値が選ばれる。

【0086】また、コードブックの学習時には、上記各歪尺度をもとにして、一般化ロイドアルゴリズム(GLA)により学習を行う。

【0087】尚、コードブックサーチ時と学習時の歪尺度は、異なる値であっても良い。

\*【0079】この第1のベクトル量子化部640<sub>1</sub>では、1フレーム毎にベクトル量子化が行われる。加算器631からの出力は、1フレーム毎に、加算器651、661を介して重み付き距離計算器653、663にそれぞれ送られ、最小となる重み付き距離が算出される。

【0080】量子化誤差 $X_2$ と量子化値 $X'_2$ との差分

は、 $10 \times 2$ の行列であり、

$$X_2 - X'_2 = [x_{3-1}, x_{3-2}]$$

と表すときの、この第1のベクトル量子化部640<sub>1</sub>のベクトル量子化器652、662によるコードブックサーチ時の歪尺度 $d_{vq1}$ 、 $d_{vq2}$ を、(4)、(5)式で示す。

【0081】

【数4】

40

【0088】上記マトリクス量子化器622、632、ベクトル量子化器652、662、672、682からの各8ビットのインデックスは、信号切換器690で切り換えられて、出力端子691から出力される。

【0089】具体的には、低ビットレート時には、上記第1のマトリクス量子化工程を行う第1のマトリクス量子化部620<sub>1</sub>、上記第2のマトリクス量子化工程を行う第2のマトリクス量子化部620<sub>2</sub>、及び上記第1のベクトル量子化工程を行う第1のベクトル量子化部640<sub>1</sub>での出力を取り出し、高ビットレート時には、上記低ビットレート時の出力に上記第2のベクトル量子化工程を行う第2のベクトル量子化部640<sub>2</sub>での出力を合わせて取り出す。

【0090】これにより、2 kbps 時には、32 bits/40 msec のインデックスが出力され、6 kbps 時には、

50

48bits/40msecのインデックスが出力される。

【0091】また、上記マトリクス量子化部620及び上記ベクトル量子化部640では、上記LPC係数を表現するパラメータの持つ特性に合わせた、周波数軸方向又は時間軸方向、あるいは周波数軸及び時間軸方向に制限を持つ重み付けを行う。

【0092】まず、LSPパラメータの持つ特性に合わせた、周波数軸方向に制限を持つ重み付けについて説明する。例えば、次数 $P=10$ とすると、LSPパラメータ $x(i)$ を、低域、中域、高域の3つの領域として、

$$L_1 = \{x(i) \mid 1 \leq i \leq 2\}$$

$$L_2 = \{x(i) \mid 3 \leq i \leq 6\}$$

$$L_3 = \{x(i) \mid 7 \leq i \leq 10\}$$

とグループ化する。そして、各グループ $L_1$ 、 $L_2$ 、 $L_3$ の重み付けを $1/4$ 、 $1/2$ 、 $1/4$ とすると、各グループ $L_1$ 、 $L_2$ 、 $L_3$ の周波数軸方向のみに制限を持つ重みは、(8)、(9)、(10)式となる。

【0093】

【数6】

$$w'(i) = \frac{w(i)}{\sum_{j=1}^2 w(j)} \times \frac{1}{4} \quad \dots(8)$$

$$w'(i) = \frac{w(i)}{\sum_{j=3}^6 w(j)} \times \frac{1}{2} \quad \dots(9)$$

$$w'(i) = \frac{w(i)}{\sum_{j=7}^{10} w(j)} \times \frac{1}{4} \quad \dots(10)$$

【0094】これにより、各LSPパラメータの重み付けは、各グループ内でのみ行われ、その重みは各グループに対する重み付けで制限される。

【0095】ここで、時間軸方向からみると、各フレームの重み付けの総和は、必ず1となるので、時間軸方向の制限は1フレーム単位である。この時間軸方向のみに制限を持つ重みは、(11)式となる。

【0096】

【数7】

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{10} \sum_{s=0}^1 w(j,s)} \quad \dots(11) \quad (1 \leq i \leq 10, 0 \leq t \leq 1)$$

【0097】この(11)式により、周波数軸方向での制限のない、フレーム番号 $t=0, 1$ の2つのフレーム間で、重み付けが行われる。この時間軸方向のみに制限を持つ重み付けは、マトリクス量子化を行う2フレーム間で行う。

【0098】また、学習時には、学習データとして用い

る全ての音声フレーム、即ち全データのフレーム数 $T$ について、(12)式により、重み付けを行う。

【0099】

【数8】

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{10} \sum_{s=0}^1 w(j,s)} \quad \dots(12) \quad (1 \leq i \leq 10, 0 \leq t \leq T)$$

【0100】また、周波数軸方向及び時間軸方向に制限を持つ重み付けについて説明する。例えば、次数 $P=10$ とすると、LSPパラメータ $x(i, t)$ を、低域、中域、高域の3つの領域として、

$$L_1 = \{x(i, t) \mid 1 \leq i \leq 2, 0 \leq t \leq 1\}$$

$$L_2 = \{x(i, t) \mid 3 \leq i \leq 6, 0 \leq t \leq 1\}$$

$$L_3 = \{x(i, t) \mid 7 \leq i \leq 10, 0 \leq t \leq 1\}$$

とグループ化する。各グループ $L_1$ 、 $L_2$ 、 $L_3$ の重み付けを $1/4$ 、 $1/2$ 、 $1/4$ とすると、各グループ $L_1$ 、 $L_2$ 、 $L_3$ の周波数軸方向及び時間軸方向に制限を持つ重み付けは、(13)、(14)、(15)式となる。

【0101】

【数9】

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^2 \sum_{s=0}^1 w(j,s)} \times \frac{1}{4} \quad \dots(13)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=3}^6 \sum_{s=0}^1 w(j,s)} \times \frac{1}{2} \quad \dots(14)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=7}^{10} \sum_{s=0}^1 w(j,s)} \times \frac{1}{4} \quad \dots(15)$$

【0102】この(13)、(14)、(15)式により、周波数軸方向では3つの帯域毎に、時間軸方向ではマトリクス量子化を行う2フレーム間に重み付けの制限を加えた重み付けを行う。これは、コードブックサーチ時及び学習時共に有効となる。

【0103】また、学習時には、全データのフレーム数について重み付けを行う。LSPパラメータ $x(i, t)$ を、低域、中域、高域の3つの領域として、 $L_1 = \{x(i, t) \mid 1 \leq i \leq 2, 0 \leq t \leq T\}$   
 $L_2 = \{x(i, t) \mid 3 \leq i \leq 6, 0 \leq t \leq T\}$   
 $L_3 = \{x(i, t) \mid 7 \leq i \leq 10, 0 \leq t \leq T\}$ とグループ化し、各グループ $L_1$ 、 $L_2$ 、 $L_3$ の重み付けを $1/4$ 、 $1/2$ 、 $1/4$ とすると、各グループ $L_1$ 、 $L_2$ 、 $L_3$ の周波数軸方向及び時間軸方向に制限を持つ重み付けは、(16)、(17)、(18)式となる。

【0104】

【数10】

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{21} \sum_{s=0}^T w(j,s)} \times \frac{1}{4} \quad \dots(16)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=3}^6 \sum_{s=0}^T w(j,s)} \times \frac{1}{2} \quad \dots(17)$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=7}^{10} \sum_{s=0}^T w(j,s)} \times \frac{1}{4} \quad \dots(18)$$

【0105】この(16)、(17)、(18)式により、周波数軸方向では3つの帯域毎に重み付けを行い、時間軸方向では全フレーム間で重み付けを行うことができる。

【0106】さらに、上記マトリクス量子化部620及び上記ベクトル量子化部640では、上記LSPパラメータの変化の大きさに応じて重み付けを行う。音声フレーム全体においては少数フレームとなる、 $V \rightarrow UV$ 、 $UV \rightarrow V$ の遷移(トランジェント)部において、子音と母音との周波数特性の違いから、LSPパラメータは大きく変化する。そこで、(19)式に示す重みを、上述の重み $w'(i, t)$ に乗算することにより、上記遷移部を重視する重み付けを行うことができる。

【0107】

【数11】

$$wd(t) = \sum_{i=1}^{10} |x_1(i,t) - x_1(i,t-1)|^2 \quad \dots(19)$$

【0108】尚、(19)式の代わりに、(20)式を用いることも考えられる。

【0109】

【数12】

$$wd(t) = \sum_{i=1}^{10} \sqrt{|x_1(i,t) - x_1(i,t-1)|} \quad \dots(20)$$

【0110】このように、LSP量子化器134では、2段のマトリクス量子化及び2段のベクトル量子化を行うことにより、出力するインデックスのビット数を可変にすることができる。

【0111】次に、上記図1、図3のベクトル量子化部116の基本構成を図7に、図7のベクトル量子化部のより具体的な構成を図8にそれぞれ示し、ベクトル量子化部116におけるスペクトルエンベロープ(Am)の重み付きベクトル量子化の具体例について説明する。

【0112】先ず、図3の音声信号符号化装置において、スペクトル評価部148の出力側あるいはベクトル量子化器116の入力側に設けられたスペクトルエンベロープの振幅のデータ数を一定個数にするデータ数変換の具体例について説明する。

【0113】このデータ数変換には種々の方法が考えられるが、本実施の形態においては、例えば、周波数軸上の有効帯域1ブロック分の振幅データに対して、ブロック内の最後のデータからブロック内の最初のデータまでの値を補間するようなダミーデータ、あるいはブロックの最後のデータ、最初のデータを繰り返すような所定のデータを付加してデータ個数を $N_0$ 個に拡大した後、帯域制限型の $O_0$ 倍(例えば8倍)のオーバーサンプリングを施すことにより $O_0$ 倍の個数の振幅データを求め、この $O_0$ 倍の個数 $(m_{ux} + 1) \times O_0$ 個の振幅データを直線補間してさらに多くの $N_u$ 個(例えば2048個)に拡張し、この $N_u$ 個のデータを間引いて上記一定個数 $M$ (例えば44個)のデータに変換している。実際には、最終的に必要な $M$ 個のデータを作成するのに必要なデータのみをオーバーサンプリング及び直線補間で算出しており、 $N_u$ 個のデータを全て求めてはいない。

【0114】図3の重み付きベクトル量子化を行うベクトル量子化器116は、図7に示すように、第1のベクトル量子化工程を行う第1のベクトル量子化部500と、この第1のベクトル量子化部500における第1のベクトル量子化の際の量子化誤差ベクトルを量子化する第2のベクトル量子化工程を行う第2のベクトル量子化部510とを少なくとも有する。この第1のベクトル量子化部500は、いわゆる1段目のベクトル量子化部であり、第2のベクトル量子化部510は、いわゆる2段目のベクトル量子化部である。

【0115】第1のベクトル量子化部500の入力端子501には、スペクトル評価部148の出力ベクトル $\underline{x}$ 、即ち一定個数 $M$ のエンベロープデータが入力される。この出力ベクトル $\underline{x}$ は、ベクトル量子化器502で重み付きベクトル量子化される。これにより、ベクトル量子化器502から出力されるシェイプインデックスは出力端子503から出力され、また、量子化値 $\underline{x}_0$ は出力端子504から出力されると共に、加算器505、513に送られる。加算器505では、ソースベクトル $\underline{x}$ から量子化値 $\underline{x}_0$ が減算されて、量子化誤差ベクトル $\underline{y}$ が得られる。

【0116】この量子化誤差ベクトル $\underline{y}$ は、第2のベクトル量子化部510内のベクトル量子化部511に送られる。このベクトル量子化部511は、複数個のベクトル量子化器で構成され、図7では、2個のベクトル量子化器511<sub>1</sub>、511<sub>2</sub>から成る。量子化誤差ベクトル $\underline{y}$ は次元分割されて、2個のベクトル量子化器511<sub>1</sub>、511<sub>2</sub>で、それぞれ重み付きベクトル量子化される。これらのベクトル量子化器511<sub>1</sub>、511<sub>2</sub>から出力されるシェイプインデックスは、出力端子512<sub>1</sub>、512<sub>2</sub>からそれぞれ出力され、また、量子化値 $\underline{y}_1$ 、 $\underline{y}_2$ は次元方向に接続されて、加算器513に送られる。この加算器513では、量子化値 $\underline{y}_1$ 、 $\underline{y}_2$ と量子化値 $\underline{x}_0$ とが加算されて、量子化値 $\underline{x}_1$ が生成される。この

量子化値  $\underline{x}_1'$  は出力端子514から出力される。

【0117】これにより、低ビットレート時には、上記第1のベクトル量子化部500による第1のベクトル量子化工程での出力を取り出し、高ビットレート時には、上記第1のベクトル量子化工程での出力及び上記第2の量子化部510による第2のベクトル量子化工程での出力を取り出す。

【0118】具体的には、図8に示すように、ベクトル量子化器116内の第1のベクトル量子化部500のベクトル量子化器502は、L次元、例えば4次元の2

【0119】すなわち、4次元でコードブックサイズが32のベクトル量子化コードブックからの出力ベクトル\*

$$E = \| \mathbf{W} \{ \mathbf{H} \underline{x} - \mathbf{H} \mathbf{g}_1 (\underline{s}_{01} + \underline{s}_{11}) \} \|^2 \quad \dots (21)$$

$$= \| \mathbf{W} \mathbf{H} \{ \underline{x} - \mathbf{g}_1 (\underline{s}_{01} + \underline{s}_{11}) \} \|^2$$

と定義する。この(21)式において、 $\mathbf{H}$ はLPCの合成フィルタの周波数軸上での特性であり、 $\mathbf{W}$ は聴覚重み付けの周波数軸上での特性を表す重み付けのための行列である。

【0122】行列 $\mathbf{H}$ は、現フレームのLPC分析結果

【0123】

【数13】

$$H(z) = \frac{1}{1 + \sum_{i=1}^P \alpha_i z^{-i}} \quad \dots (22)$$

【0124】の周波数特性からL次元、例えば4次元の各対応する点の値をサンプルしたものである。

【0125】算出手順としては、一例として、1、 $\alpha_1, \alpha_2, \dots, \alpha_P$ に0詰めして、すなわち、1、 $\alpha_1, \alpha_2, \dots, \alpha_P, 0, 0, \dots, 0$ として、例えば256点のデータにする。その後、256点FFTを行い、 $(\text{re}^2 + \text{im}^2)^{1/2}$ を $0 \sim \pi$ に対応する点に対して算出して、その逆数をとる。それをL点、すなわち例えば44点に間引いたものを対角要素とする行列を、

【0126】

【数14】

$$\mathbf{H} = \begin{bmatrix} h(1) & & 0 \\ & h(2) & \\ 0 & & h(L) \end{bmatrix}$$

【0127】とする。

【0128】聴覚重み付け行列 $\mathbf{W}$ は、以下のように求められる。

【0129】

【数15】

\*ルの和に、ゲイン $g_1$ を乗じたものを、4次元のスペクトルエンベロープベクトル $\underline{x}$ の量子化値 $\underline{x}_0'$ として使用する。これは、図8に示すように、2つのシェイプコードブックをCB0、CB1とし、その出力ベクトルを $\underline{s}_{01}, \underline{s}_{11}$ 、ただし $0 \leq i, j \leq 31$ 、とする。また、ゲインコードブックCBgの出力を $g_1$ 、ただし $0 \leq 1 \leq 31$ 、とする。 $g_1$ はスカラー値である。この最終出力 $\underline{x}_0'$ は、 $g_1 (\underline{s}_{01} + \underline{s}_{11})$ となる。

【0120】LPC残差について上記MBE分析によって得られたスペクトルエンベロープAmを一定次元に変換したものを $\underline{x}$ とする。このとき、 $\underline{x}$ をいかに効率的に量子化するかが重要である。

【0121】ここで、量子化誤差エネルギーEを、

$$W(z) = \frac{1 + \sum_{i=1}^P \alpha_i \lambda_b^i z^{-i}}{1 + \sum_{i=1}^P \alpha_i \lambda_a^i z^{-i}} \quad \dots (23)$$

【0130】この(23)式で、 $\alpha_i$ は入力LPC分析結果である。また、 $\lambda_a, \lambda_b$ は定数であり、一例として、 $\lambda_a = 0.4, \lambda_b = 0.9$ が挙げられる。

【0131】行列あるいはマトリクス $\mathbf{W}$ は、上記(23)式の周波数特性から算出できる。一例として、1、 $\alpha_1 \lambda_b, \alpha_2 \lambda_b^2, \dots, \alpha_P \lambda_b^P, 0, 0, \dots, 0$ として256点のデータとしてFFTを行い、 $0$ 以上 $\pi$ 以下の区間に対して $(\text{re}^2 + \text{im}^2)^{1/2}, 0 \leq i \leq 128$ 、を求める。次に、1、 $\alpha_1 \lambda_a, \alpha_2 \lambda_a^2, \dots, \alpha_P \lambda_a^P, 0, 0, \dots, 0$ として分母の周波数特性を256点FFTで $0 \sim \pi$ の区間を128点で算出する。これを $(\text{re}'^2 + \text{im}'^2)^{1/2}, 0 \leq i \leq 128$ 、とする。

【0132】

【数16】

$$w_0[i] = \frac{\sqrt{\text{re}^2[i] + \text{im}^2[i]}}{\sqrt{\text{re}'^2[i] + \text{im}'^2[i]}} \quad (0 \leq i \leq 128)$$

【0133】として、上記(23)式の周波数特性が求められる。

【0134】これをL次元、例えば4次元ベクトルの対応する点について、以下の方法で求める。より正確には、直線補間を用いるべきであるが、以下の例では最も近い点の値で代用している。

【0135】すなわち、

$$\omega[i] = \omega_0[\text{nint}(128i/L)] \quad 1 \leq i \leq L$$

ただし、 $\text{nint}(X)$ は、Xに最も近い整数を返す関数である。

【0136】また、上記 $\mathbf{H}$ に関しても同様の方法で、 $h(1), h(2), \dots, h(L)$ を求めている。すなわち、

【0137】

$$H = \begin{bmatrix} h(1) & & 0 \\ & h(2) & \\ 0 & & h(L) \end{bmatrix} \quad W = \begin{bmatrix} w(1) & & 0 \\ & w(2) & \\ 0 & & w(L) \end{bmatrix}$$

$$WH = \begin{bmatrix} h(1)w(1) & & 0 \\ & h(2)w(2) & \\ 0 & & h(L)w(L) \end{bmatrix} \quad \dots (24)$$

【0138】となる。

【0139】ここで、他の例として、FFTの回数を減らすのに、 $H(z)W(z)$ を先に求めてから、周波数特性を求めてもよい。すなわち、

【0140】

【数18】

$$H(z)W(z) = \frac{1}{1 + \sum_{i=1}^P \alpha_i z^{-i}} \cdot \frac{1 + \sum_{i=1}^P \alpha_i \lambda_a^i z^{-i}}{1 + \sum_{i=1}^P \alpha_i \lambda_a^i z^{-i}} \quad \dots (25)$$

【0141】この(25)式の分母を展開した結果を、

【0142】

【数19】

$$\left(1 + \sum_{i=1}^P \alpha_i z^{-i}\right) \left(1 + \sum_{i=1}^P \alpha_i \lambda_a^i z^{-i}\right) = 1 + \sum_{i=1}^{2P} \beta_i z^{-i}$$

【0143】とする。ここで、 $1, \beta_1, \beta_2, \dots, \beta_{2P}, 0, 0, \dots, 0$ として、例えば256点のデータにする。その後、256点FFTを行い、振幅の周

波数特性を、

【0144】

【数20】

$$rms[i] = \sqrt{re^{*2}[i] + im^{*2}[i]} \quad 0 \leq i \leq 128$$

【0145】とする。これより、

【0146】

【数21】

$$wh_0[i] = \frac{\sqrt{re^2[i] + im^2[i]}}{\sqrt{re^{*2}[i] + im^{*2}[i]}} \quad 0 \leq i \leq 128$$

【0147】これをL次元ベクトルの対応する点について求める。上記FFTのポイント数が少ない場合は、直線補間で求めるべきであるが、ここでは最寄りの値を使用している。すなわち、

【0148】

【数22】

$$wh[i] = wh_0 \left[ \text{nint} \left( \frac{128}{L} \cdot i \right) \right] \quad 1 \leq i \leq L$$

【0149】である。これを対角要素とする行列を  $W'$  とすると、

【0150】

【数23】

$$W' = \begin{bmatrix} wh(1) & & 0 \\ & wh(2) & \\ 0 & & wh(L) \end{bmatrix} \quad \dots (26)$$

【0151】となる。(26)式は上記(24)式と同一のマトリクスとなる。

【0152】あるいは、(25)式より直接  $|H(\exp(j\omega))W(\exp(j\omega))|$  を  $\omega = i\pi/L$  (ただし、 $1 \leq i \leq L$ ) に関して算出したものを  $wh[i]$  に使用してもよい。又は、(25)式のインパルス応答を適当な長さ(例えば40点)求めて、それを用いてFFTして振幅周波数特性を求めて使用してもよい。

【0153】このマトリクス、すなわち重み付き合成フィルタの周波数特性を用いて、上記(21)式を書き直すと、

【0154】

【数24】

$$E = \|W'(\underline{x} - g_i(\underline{s}_{0i} + \underline{s}_{1j}))\| \quad \dots (27)$$

【0155】となる。

【0156】ここで、シェイプコードブックとゲインコードブックの学習法について説明する。

【0157】先ず、CBOに関しコードベクトル  $\underline{s}_{0c}$  を選択する全てのフレーム  $k$  に関して歪の期待値を最小化する。そのようなフレームが  $M$  個あるとして、

【0158】

【数25】

$$J = \frac{1}{M} \sum_{k=1}^M \|W_k(\underline{x}_k - g_k(\underline{s}_{0c} + \underline{s}_{1k}))\|^2 \quad \dots (28)$$

【0159】を最小化すればよい。この(28)式中



で、 $\mathbf{W}_k$  はk番目のフレームに対する重み、 $\mathbf{x}_k$  はk番目のフレームの入力、 $g_k$  はk番目のフレームのゲイン、 $\mathbf{s}_{1k}$  はk番目のフレームについてのコードブックC B1からの出力、をそれぞれ示す。

\*

$$\begin{aligned} J &= \frac{1}{M} \sum_{k=1}^M \left\{ \left( \mathbf{x}_k^T - g_k (\mathbf{s}_{0c}^T + \mathbf{s}_{1k}^T) \right) \mathbf{W}_k'^T \mathbf{W}_k' \left( \mathbf{x}_k - g_k (\mathbf{s}_{0c} + \mathbf{s}_{1k}) \right) \right\} \\ &= \frac{1}{M} \sum_{k=1}^M \left\{ \mathbf{x}_k^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{x}_k - 2g_k (\mathbf{s}_{0c}^T + \mathbf{s}_{1k}^T) \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{x}_k \right. \\ &\quad \left. + g_k^2 (\mathbf{s}_{0c}^T + \mathbf{s}_{1k}^T) \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0c} + \mathbf{s}_{1k}) \right\} \\ &= \frac{1}{M} \sum_{k=1}^M \left\{ \mathbf{x}_k^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{x}_k - 2g_k (\mathbf{s}_{0c}^T + \mathbf{s}_{1k}^T) \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{x}_k \right. \\ &\quad \left. + g_k^2 \mathbf{s}_{0c}^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}_{0c} + 2g_k^2 \mathbf{s}_{0c}^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}_{1k} \right. \\ &\quad \left. + g_k^2 \mathbf{s}_{1k}^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}_{1k} \right\} \quad \cdots (29) \end{aligned}$$

【0162】

※ ※ 【数27】

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{s}_{0c}} &= \frac{1}{M} \sum_{k=1}^M \left\{ -2g_k \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{x}_k + 2g_k^2 \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}_{0c} \right. \\ &\quad \left. + 2g_k^2 \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}_{1k} \right\} = 0 \quad \cdots (30) \\ \text{より} \quad \sum_{k=1}^M \left( g_k \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{x}_k - g_k^2 \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}_{1k} \right) &= \sum_{k=1}^M g_k^2 \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}_{0c} \end{aligned}$$

よって

$$\mathbf{s}_{0c} = \left\{ \sum_{k=1}^M g_k^2 \mathbf{W}_k'^T \mathbf{W}_k' \right\}^{-1} \cdot \left\{ \sum_{k=1}^M g_k^2 \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{x}_k - g_k \mathbf{s}_{1k}) \right\} \quad \cdots (31)$$

{ }<sup>-1</sup>は逆行列を意味する。 $\mathbf{W}_k'^T$ は $\mathbf{W}_k'$ の転置行列である。

【0163】次に、ゲインに関する最適化を考える。 ★【0165】

【0164】ゲインのコードワード $g_c$ を選択するk番

【数28】

目のフレームに関する歪の期待値 $J_g$ は、

★30

$$\begin{aligned} J_g &= \frac{1}{M} \sum_{k=1}^M \left\| \mathbf{W}_k' (\mathbf{x}_k - g_c (\mathbf{s}_{0k} + \mathbf{s}_{1k})) \right\|^2 \\ &= \frac{1}{M} \sum_{k=1}^M \left\{ \mathbf{x}_k^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{x}_k - 2g_c \mathbf{x}_k^T \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k}) \right. \\ &\quad \left. + g_c^2 (\mathbf{s}_{0k}^T + \mathbf{s}_{1k}^T) \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k}) \right\} \\ \frac{\partial J_g}{\partial g_c} &= \frac{1}{M} \sum_{k=1}^M \left\{ -2\mathbf{x}_k^T \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k}) \right. \\ &\quad \left. + 2g_c (\mathbf{s}_{0k}^T + \mathbf{s}_{1k}^T) \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k}) \right\} = 0 \text{を解いて} \\ \sum_{k=1}^M \mathbf{x}_k^T \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k}) &= \sum_{k=1}^M g_c (\mathbf{s}_{0k}^T + \mathbf{s}_{1k}^T) \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k}) \\ g_c &= \frac{\sum_{k=1}^M \mathbf{x}_k^T \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k})}{\sum_{k=1}^M (\mathbf{s}_{0k}^T + \mathbf{s}_{1k}^T) \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{s}_{0k} + \mathbf{s}_{1k})} \quad \cdots (32) \end{aligned}$$

【0166】上記(31)式及び(32)式は、シェイプ $\mathbf{s}_{0i}$ 、 $\mathbf{s}_{1i}$ 及びゲイン $g_i$ 、 $0 \leq i \leq 31$ 、 $0 \leq j \leq 31$ 、 $0 \leq l \leq 31$ の最適なセントロイドコンディション(Centroid Condition)、すなわち最適なデコーダ出力

を与えるものである。なお、 $\mathbf{s}_{1i}$ に関しても $\mathbf{s}_{0i}$ と同様に求めることができる。

【0167】次に、最適エンコード条件(Nearest Neighbour Condition)を考える。

【0168】歪尺度を求める上記(27)式、すなわち、  
 $E = \| \mathbf{W}' (\mathbf{x} - \mathbf{g}_i (\mathbf{s}_{0i} + \mathbf{s}_{1i})) \|^2$   
 を最小化する  $\mathbf{s}_{0i}$ 、 $\mathbf{s}_{1i}$  を、入力  $\mathbf{x}$ 、重みマトリクス  $\mathbf{W}'$  が与えられる毎に、すなわち毎フレームごとに決定する。

【0169】このようなコードブックサーチは、本来は、総当りの全ての  $\mathbf{g}_i$  ( $0 \leq i \leq 31$ )、 $\mathbf{s}_{0i}$  ( $0 \leq i \leq 31$ )、 $\mathbf{s}_{1i}$  ( $0 \leq j \leq 31$ ) の組み合わせの、 $32 \times 32 \times 32 = 32768$  通りについて  $E$  を求めて、10 最小の  $E$  を与える  $\mathbf{g}_i$ 、 $\mathbf{s}_{0i}$ 、 $\mathbf{s}_{1i}$  の組を求めるべきで \*

$$E = \|\mathbf{x}_w - \mathbf{g}_i \mathbf{s}_w\|^2 \quad \dots (33)$$

$$E = \|\mathbf{s}_w\|^2 + \|\mathbf{s}_w\|^2 \left( \mathbf{g}_i - \frac{\mathbf{x}_w^T \cdot \mathbf{s}_w}{\|\mathbf{s}_w\|^2} \right)^2 - \frac{(\mathbf{x}_w^T \cdot \mathbf{s}_w)^2}{\|\mathbf{s}_w\|^2} \quad \dots (34)$$

【0172】となる。従って、 $\mathbf{g}_i$  の精度が充分にとれ  
 ると仮定すると、

【0173】

【数30】

- (1)  $\frac{(\mathbf{x}_w^T \cdot \mathbf{s}_w)^2}{\|\mathbf{s}_w\|^2}$  の最大となる  $\mathbf{s}_w$  をサーチ  
 (2)  $\frac{\mathbf{x}_w^T \cdot \mathbf{s}_w}{\|\mathbf{s}_w\|^2}$  にもっとも近い  $\mathbf{g}_i$  をサーチ

【0174】という2つのステップに分けてサーチする  
 ことができる。元の表記を用いて書き直すと、

【0175】

【数31】

- (1)'  $\frac{(\mathbf{x}^T \mathbf{W}'^T \mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1j}))^2}{\|\mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1j})\|^2}$  を最大化する  
 $\mathbf{s}_{0i} \mathbf{s}_{1j}$  の組をサーチ  
 (2)'  $\frac{(\mathbf{x}^T \mathbf{W}'^T \mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1j}))^2}{\|\mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1j})\|^2}$   
 にもっとも近い  $\mathbf{g}_i$  をサーチ  
 ... (35)

【0176】となる。この(35)式が最適エンコード  
 条件(Nearest Neighbour Condition)である。

【0177】次に、このようなベクトル量子化のコード  
 ブックサーチ(符号帳検索)を行う場合の演算量につ  
 てさらに考察する。

【0178】先ず、上記(35)式の(1)'の演算量 50

\*あるが、膨大な演算量となるので、本実施の形態では、  
 シェイプとゲインのシーケンシャルサーチを行っている。  
 なお、 $\mathbf{s}_{0i}$  と  $\mathbf{s}_{1i}$  との組み合わせについては、総当り  
 サーチを行うものとする。これは、 $32 \times 32 = 1024$   
 通りである。以下の説明では、簡単化のため、 $\mathbf{s}_{0i} + \mathbf{s}_{1i}$   
 を  $\mathbf{s}_w$  と記す。

【0170】上記(27)式は、 $E = \| \mathbf{W}' (\mathbf{x} - \mathbf{g}_i \mathbf{s}_w) \|^2$  となる。さらに簡単のため、 $\mathbf{x}_w = \mathbf{W}' \mathbf{x}$ 、  
 $\mathbf{s}_w = \mathbf{W}' \mathbf{s}_w$  とすると、

【0171】

【数29】

は、 $\mathbf{s}_{0i}$  及び  $\mathbf{s}_{1i}$  の次元を  $K$ 、コードブック  $\mathbf{C}B_0$ 、 $\mathbf{C}$   
 20  $B_1$  のサイズをそれぞれ  $L_0$ 、 $L_1$ 、すなわち

$0 \leq i < L_0$ 、 $0 \leq j < L_1$

とし、分子の加算、積和、2乗の各演算量をそれぞれ  
 1、分母の積、積和の各演算量をそれぞれ1として、概  
 略、

分子 :  $L_0 \cdot L_1 \cdot \{K \cdot (1+1) + 1\}$

分母 :  $L_0 \cdot L_1 \cdot K \cdot (1+1)$

大小比較 :  $L_0 \cdot L_1$

となり、計  $L_0 \cdot L_1 (4K+2)$  となる。ここで、 $L_0$   
 $= L_1 = 32$ 、 $K = 44$  とすると、演算量は、1822

72 程度のオーダーとなる。

【0179】そこで、上記(35)式の(1)'の演算  
 を全て実行せずに、 $\mathbf{s}_{0i}$  及び  $\mathbf{s}_{1i}$  に関して  $P$  個ずつ予備  
 選択(プリセクション)を行う。なお、ここでは、負  
 のゲインエントリを考慮していない(許していない)た  
 め、上記(35)式の(2)'の分子の値は常に正の数  
 となるように、上記(35)式の(1)'のサーチを行  
 う。すなわち、 $\mathbf{x}^T \mathbf{W}'^T \mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1i})$  の極性  
 込みで、上記(35)式の(1)'の最大化を行う。

【0180】このような予備選択方法の具体例につ  
 40 いて説明すると、先ず、

(手順1)  $\mathbf{x}^T \mathbf{W}'^T \mathbf{W}' \mathbf{s}_{0i}$  を最大にする  $\mathbf{s}_{0i}$  を上位  
 から  $P_0$  個選択

(手順2)  $\mathbf{x}^T \mathbf{W}'^T \mathbf{W}' \mathbf{s}_{1i}$  を最大にする  $\mathbf{s}_{1i}$  を上位  
 から  $P_1$  個選択

(手順3) これら  $P_0$  個の  $\mathbf{s}_{0i}$  と  $P_1$  個の  $\mathbf{s}_{1i}$  の全ての組  
 組み合わせについて、上記(35)式の(1)'の式を評  
 価

という方法が挙げられる。

【0181】これは、上記(35)式の(1)'の式の  
 平方根である、

【0182】

【数32】

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'(\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}'(\mathbf{s}_{0i} + \mathbf{s}_{1j})\|} \dots (a1)$$

【0183】の評価において、分母、すなわち  $\mathbf{s}_{0i} + \mathbf{s}_{1j}$  の重み付きノルムが、 $i, j$  によらずほぼ一定という仮定が成立するときに有効である。実際には上記 (a1) 式の分母の大きさは一定ではないが、これを考慮した予備選択方法については、後述する。

【0184】ここでは、上記 (a1) 式の分母が一定と仮定した場合の演算量の削減効果を説明する。上記 (手順1) のサーチに  $L_0 \cdot K$  の演算量を要し、大小比較に、

$$(L_0 - 1) + (L_0 - 2) + \dots + (L_0 - P_0) \\ = P_0 \cdot L_0 - P_0 \cdot (1 + P_0) / 2$$

を要するから、演算量の計は、 $L_0 \cdot (K + P_0) - P_0 \cdot (1 + P_0) / 2$  となる。また、上記 (手順2) にも同様の処理量が必要であり、これらを合計して、予備選択のための演算処理量は、

$$L_0 \cdot (K + P_0) + L_1 \cdot (K + P_1) - P_0 \cdot (1 + P_0) / 2 - P_1 \cdot (1 + P_1) / 2$$

となる。

【0185】また、上記 (手順3) の本選択の処理については、上記 (35) 式の (1)' の演算に関して、

分子 :  $P_0 \cdot P_1 \cdot (1 + K + 1)$ 分母 :  $P_0 \cdot P_1 \cdot K \cdot (1 + 1)$ 大小比較 :  $P_0 \cdot P_1$ となり、計  $P_0 \cdot P_1 \cdot (3K + 3)$  となる。

【0186】例えば、 $P_0 = P_1 = 6$ 、 $L_0 = L_1 = 32$ 、 $K = 44$  とすると、演算量は、本選択で 4860、予備選択で 3158 となり、計 8018 程度のオーダーとなる。また、予備選択の個数をそれぞれ 10 個にまで増やして  $P_0 = P_1 = 10$  としても、本選択で 13500、予備選択で 3346 となり、16846 程度のオーダーとなる。

【0187】このように、予備選択するベクトルの個数を各コードブックそれぞれ 10 個ずつとした場合でも、前述した全てを演算する場合の 182272 と比較して、

$$16846 / 182272$$

となり、元の約 1/10 以下の演算量に抑えることができる。

【0188】ところで、上記 (35) 式の (1)' の式の分母の大きさは一定ではなく、選択されたコードベクトルに依存して大小変化する。そこで、ある程度このノルムの概略の大きさを考慮した予備選択 (プリセクション) 方法について、以下に説明する。

【0189】上記 (35) 式の (1)' の式の平方根である上記 (a1) 式の最大値を求める場合に、

【0190】

【数33】

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'(\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|} \leq \frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'(\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}'(\mathbf{s}_{0i} + \mathbf{s}_{1j})\|} \dots (a2)$$

【0191】であることを考慮して、この (a2) 式の左辺を最大化すればよい。そこで、この左辺を

【0192】

【数34】

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'(\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|} \\ = \frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'\mathbf{s}_{0i}}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|} + \frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'\mathbf{s}_{1j}}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|} \dots (a3)$$

【0193】のように展開して、この (a3) 式の第1項、第2項をそれぞれ最大化する。

【0194】上記 (a3) 式の第1項の分子は  $\mathbf{s}_{0i}$  のみの関数なので、 $\mathbf{s}_{0i}$  に関する最大化を考える。また上記 (a3) 式の第2項の分子は  $\mathbf{s}_{1j}$  のみの関数なので、 $\mathbf{s}_{1j}$  に関する最大化を考える。すなわち、

【0195】

【数35】

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'\mathbf{s}_{0i}}{\|\mathbf{W}'\mathbf{s}_{0i}\|} \dots (a4)$$

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}'\mathbf{s}_{1j}}{\|\mathbf{W}'\mathbf{s}_{1j}\|} \dots (a5)$$

【0196】において、

(手順1) 上記 (a4) 式を最大化するものの上位から  $Q_0$  個の  $\mathbf{s}_{0i}$  を選択

(手順2) 上記 (a5) 式を最大化するものの上位から  $Q_1$  個の  $\mathbf{s}_{1j}$  を選択

(手順3) 選択された  $Q_0$  個の  $\mathbf{s}_{0i}$  と  $Q_1$  個の  $\mathbf{s}_{1j}$  の全ての組み合わせについて、上記 (35) 式の (1)' の式を評価

という方法が挙げられる。

【0197】なお、 $\mathbf{W}' = \mathbf{W}\mathbf{H} / \|\mathbf{x}\|$  であり、 $\mathbf{W}$  も  $\mathbf{H}$  も入力ベクトル  $\mathbf{x}$  の関数であり、当然  $\mathbf{W}'$  も入力ベクトル  $\mathbf{x}$  の関数となる。

【0198】従って、本来は入力ベクトル  $\mathbf{x}$  毎に  $\mathbf{W}'$  を算出し、上記 (a4)、(a5) 式の分母を計算すべきであるが、ここでは予備選択なのであまり演算処理量を費やしたくない。そこで、これらの分母については、典型的な、すなわち代表的な  $\mathbf{W}'$  の値を用いて事前に各  $\mathbf{s}_{0i}$ 、 $\mathbf{s}_{1j}$  について計算した値を、 $\mathbf{s}_{0i}$ 、 $\mathbf{s}_{1j}$  と共にテーブルに格納しておく。なお、実際のサーチ演算で割り算をすることは処理が重いので、

【0199】

【数36】

$$\frac{1}{\|\mathbf{W}_{s_{0i}}^*\|} \quad 0 \leq i < L_0 \quad \cdots (a6)$$

$$\frac{1}{\|\mathbf{W}_{s_{1j}}^*\|} \quad 0 \leq j < L_1 \quad \cdots (a7)$$

$$\mathbf{W}^* = \frac{1}{N} \sum_{k=1}^N \mathbf{W}'_k \quad \cdots (a8)$$

ただし、 $\mathbf{W}'_k$ はV/UVが有声音と判定されたフレームの

$\mathbf{W}'$ で

$$\mathbf{W}' = \frac{\mathbf{W}\mathbf{H}}{\|\mathbf{x}\|} \quad \cdots (a9)$$

【0202】図9は、上記 $\mathbf{W}^*$ を下記の(a10)式で記述したときの各 $\mathbf{W}[0] \sim \mathbf{W}[43]$ の具体例を示すものである。

【0203】

【数38】

$$\mathbf{W}^* = \begin{bmatrix} \mathbf{W}[0] & & & 0 \\ & \mathbf{W}[1] & & \\ & & \mathbf{W}[2] & \\ & & & \ddots \\ 0 & & & & \mathbf{W}[43] \end{bmatrix} \quad \cdots (a10)$$

【0204】上記(a4)、(a5)式の分子については、 $\mathbf{W}'$ を入力ベクトル $\mathbf{x}$ 毎に算出して使用する。これは、いずれにせよ $s_{0i}$ 、 $s_{1j}$ と $\mathbf{x}$ との内積を算出せねばならないので、 $\mathbf{x}^T \mathbf{W}'^T \mathbf{W}'$ を一度計算しておけば演算量増加はごく僅かであることを考慮したものである。

【0205】このような予備選択方法に要する演算量の概略を見積もると、上記(手順1)のサーチでは、 $L_0$ 、 $(K+1)$ の演算量を要し、大小比較に、 $Q_0 \cdot L_0 = Q_0 \cdot (1 + Q_0) / 2$ を要する。また、上記(手順2)にも同様の処理量が必要であり、これらを合計して、予備選択のための演算処理量は、

$$L_0(K+Q_0+1) + L_1(K+Q_1+1) - Q_0(1+Q_0)/2 - Q_1(1+Q_1)/2$$

となる。

【0206】また、上記(手順3)の本選択の処理については、上記(35)式の(1)'の演算に関して、

$$\text{分子} : Q_0 \cdot Q_1 \cdot (1 + K + 1)$$

$$\text{分母} : Q_0 \cdot Q_1 \cdot K \cdot (1 + 1)$$

$$\text{大小比較} : Q_0 \cdot Q_1$$

となり、計 $Q_0 \cdot Q_1 (3K+3)$ となる。

【0207】例えば、 $Q_0=Q_1=6$ 、 $L_0=L_1=32$ 、 $K=44$ とすると、演算量は、本選択で4860、予備選択で3222となり、計8082程度のオーダーとなる。また、予備選択の個数をそれぞれ10個にまで増やして $Q_0=Q_1=10$ としても、本選択で13500、予

\*【0200】の(a6)、(a7)式の値を格納しておく。ここで、 $\mathbf{W}^*$ は、次の(a8)式に示すものである。

【0201】

\*【数37】

備選択で3410となり、16910程度のオーダーとなる。

【0208】これらの計算結果は、上述した重み付きノルムで割り込まない(ノーマライズ無し)場合の $P_0=P_1=6$ としたときの演算量の、計8018程度のオーダーや、 $P_0=P_1=10$ としたときの16846程度のオーダーとはほぼ同程度の値であり、予備選択するベクトルの個数を各コードブックそれぞれ10個ずつとした場合でも、前述した全てを演算する場合の182272と比較して、

$$16910 / 182272$$

となり、元の約1/10以下の演算量に抑えることができる。

【0209】以上説明したような予備選択を行わないで分析・合成した音声基準として、予備選択を行った場合のSNR(SN比)と、20msecセグメントのときのセグメンタルSNRとについて具体例を示すと、ノーマライズ無し $P_0=P_1=6$ のときのSNR:14.8(dB)、セグメンタルSNR:17.5(dB)に比べて、同じ予備選択個数でノーマライズ有り、重み無しの場合、SNR:16.8(dB)、セグメンタルSNR:18.7(dB)となり、重み付きノーマライズ有りの場合、SNR:17.8(dB)、セグメンタルSNR:19.6(dB)となる。このように、ノーマライズ無しから、重み付きノーマライズ有りとすることで、2~3dBだけSNR、セグメンタルSNRが向上している。

【0210】ここで上記(31)、(32)式の条件(Centroid Condition)と、(35)式の条件を用いて、LBG(Linde-Buzo-Gray)アルゴリズム、いわゆる一般化ロイドアルゴリズム(Generalized Lloyd Algorithm:GLA)によりコードブック(CB0、CB1、CBg)を同時にトレーニングできる。

【0211】なお、本実施の形態では、 $\mathbf{W}'$ として、入力 $\mathbf{x}$ のノルムで割り込んだ $\mathbf{W}'$ を使用している。すなわち、上記(31)、(32)、(35)式において、事前に $\mathbf{W}'$ に $\mathbf{W}' / \|\mathbf{x}\|$ を代入して使用して

いる。

【0212】あるいは別法として、ベクトル量子化器116でのベクトル量子化の際の聴覚重み付けに用いられる重み $\mathbf{W}'$ については、上記(26)式で定義されているが、過去の $\mathbf{W}'$ も加味して現在の $\mathbf{W}'$ を求めることにより、テンポラルマスキングも考慮した $\mathbf{W}'$ を求めてもよい。

【0213】上記(26)式中の $wh(1), wh(2), \dots, wh(L)$

$$A_n(i) = \begin{cases} \lambda A_{n-1}(i) + (1-\lambda) wh_n(i) & (wh_n(i) \leq A_{n-1}(i)) \\ wh_n(i) & (wh_n(i) > A_{n-1}(i)) \end{cases}$$

【0216】とする。ここで、 $\lambda$ は例えば $\lambda=0.2$ とすればよい。このようにして求められた $A_n(i)$ 、 $1 \leq i \leq L$ について、これを対角要素とするマトリクスを上記重みとして用いられたい。

【0217】このように重み付きベクトル量子化により得られたシェイプインデックス $s_{01}, s_{11}$ は、出力端子520、522からそれぞれ出力され、ゲインインデックス $g_1$ は、出力端子521から出力される。また、量子化値 $x_1$ は、出力端子504から出力されると共に、加算器505に送られる。

【0218】この加算器505では、スペクトルエンベロープベクトル $\mathbf{x}$ から量子化値 $x_1$ が減算されて、量子化誤差ベクトル $\mathbf{y}$ が生成される。この量子化誤差ベクトル $\mathbf{y}$ は、具体的には、8個のベクトル量子化器511<sub>1</sub>～511<sub>8</sub>から成るベクトル量子化部511に送られて、次元分割され、各ベクトル量子化器511<sub>1</sub>～511<sub>8</sub>※

※ $h(L)$ に関して、時刻 $n$ 、すなわち第 $n$ フレームで算出されたものをそれぞれ $wh_n(1), wh_n(2), \dots, wh_n(L)$ とする。

【0214】時刻 $n$ で過去の値を考慮した重みを $A_n(i)$ 、 $1 \leq i \leq L$ と定義すると、

【0215】

【数39】

※1。で重み付きのベクトル量子化が施される。

【0219】第2のベクトル量子化部510では、第1のベクトル量子化部500と比較して、かなり多くのビット数を用いるため、コードブックのメモリ容量及びコードブックサーチのための演算量(Complexity)が非常に大きくなり、第1のベクトル量子化部500と同じ44次元のままでベクトル量子化を行うことは、不可能である。そこで、第2のベクトル量子化部510内のベクトル量子化部511を複数個のベクトル量子化器で構成し、入力される量子化値を次元分割して、複数個の低次元ベクトルとして、重み付きのベクトル量子化を行う。

【0220】ベクトル量子化器511<sub>1</sub>～511<sub>8</sub>で用いる各量子化値 $y_0 \sim y_7$ と、次元数と、ビット数との関係を表2に示す。

【0221】

【表2】

量子化値	次元	ビット数(bits)
$y_0$	4	10
$y_1$	4	10
$y_2$	4	10
$y_3$	4	10
$y_4$	4	9
$y_5$	8	8
$y_6$	8	8
$y_7$	8	7

【0222】ベクトル量子化器511<sub>1</sub>～511<sub>8</sub>から出力されるインデックス $ldvq_0 \sim ldvq_7$ は、各出力端子523<sub>1</sub>～523<sub>8</sub>からそれぞれ出力される。これらのインデックスの合計は72ビットである。

【0223】また、ベクトル量子化器511<sub>1</sub>～511<sub>8</sub>から出力される量子化値 $y_0' \sim y_7'$ を次元方向に接続した値を $\mathbf{y}'$ とすると、加算器513では、量子化値 $\mathbf{y}'$ と量子化値 $\mathbf{x}_0'$ とが加算されて、量子化値 $\mathbf{x}_1'$ が得られる。よって、この量子化値 $\mathbf{x}_1'$ は、 $\mathbf{x}_1' = \mathbf{x}_0' + \mathbf{y}'$

$$= \mathbf{x} - \mathbf{y} + \mathbf{y}'$$

で表される。すなわち、最終的な量子化誤差ベクトルは、 $\mathbf{y}' - \mathbf{y}$ となる。

【0224】尚、音声信号復号化装置側では、この第2のベクトル量子化部510からの量子化値 $\mathbf{x}_1'$ を復号化するときには、第1のベクトル量子化部500からの量子化値 $\mathbf{x}_0'$ は不要であるが、第1のベクトル量子化部500及び第2のベクトル量子化部510からのインデックスは必要とする。

50 【0225】次に、上記ベクトル量子化部511におけ

る学習法及びコードブックサーチについて説明する。

【0226】 先ず、学習法においては、量子化誤差ベクトル  $\underline{y}$  及び重み  $\underline{w}'$  を用い、表2に示すように、8つの低次元ベクトル  $\underline{y}_0 \sim \underline{y}_7$  及びマトリクスに分割する。このとき、重み  $\underline{W}'$  は、例えば44点に間引いたものを対角要素とする行列、

【0227】

【数40】

$$\underline{W}' = \begin{bmatrix} wh(1) & & & 0 \\ & wh(2) & & \\ & & \ddots & \\ 0 & & & wh(44) \end{bmatrix} \dots (36)$$

【0228】 すると、以下の8つの行列に分割される。

【0229】

【数41】

$$\begin{aligned} \underline{W}'_1 &= \begin{bmatrix} wh(1) & & 0 \\ & \ddots & \\ 0 & & wh(4) \end{bmatrix} \\ \underline{W}'_2 &= \begin{bmatrix} wh(5) & & 0 \\ & \ddots & \\ 0 & & wh(8) \end{bmatrix} \\ \underline{W}'_3 &= \begin{bmatrix} wh(9) & & 0 \\ & \ddots & \\ 0 & & wh(12) \end{bmatrix} \\ \underline{W}'_4 &= \begin{bmatrix} wh(13) & & 0 \\ & \ddots & \\ 0 & & wh(16) \end{bmatrix} \\ \underline{W}'_5 &= \begin{bmatrix} wh(17) & & 0 \\ & \ddots & \\ 0 & & wh(20) \end{bmatrix} \\ \underline{W}'_6 &= \begin{bmatrix} wh(21) & & 0 \\ & \ddots & \\ 0 & & wh(28) \end{bmatrix} \\ \underline{W}'_7 &= \begin{bmatrix} wh(29) & & 0 \\ & \ddots & \\ 0 & & wh(36) \end{bmatrix} \\ \underline{W}'_8 &= \begin{bmatrix} wh(37) & & 0 \\ & \ddots & \\ 0 & & wh(44) \end{bmatrix} \end{aligned}$$

【0230】 このように、 $\underline{y}$  及び  $\underline{W}'$  の低次元に分割されたものを、それぞれ

$\underline{y}_i, \underline{W}'_i \quad (1 \leq i \leq 8)$

とする。

【0231】 ここで、歪尺度  $E$  を、

$$E = \|\underline{W}'_i (\underline{y}_i - \underline{s})\|^2 \dots (37)$$

と定義する。このコードベクトル  $\underline{s}$  は  $\underline{y}_i$  の量子化結果であり、歪尺度  $E$  を最小化する、コードブックのコードベクトル  $\underline{s}$  がサーチされる。

10 【0232】 尚、 $\underline{W}'_i$  は、学習時には重み付けがあり、サーチ時には重み付け無し、すなわち単位行列とし、学習時とコードブックサーチ時とは異なる値を用いるようにしてもよい。

【0233】 また、コードブックの学習では、一般化ロイドアルゴリズム (GLA) を用い、さらに重み付けを行っている。先ず、学習のための最適なセントロイドコンディションについて説明する。コードベクトル  $\underline{s}$  を最適な量子化結果として選択した入力ベクトル  $\underline{y}$  が  $M$  個ある場合に、トレーニングデータを  $\underline{y}_k$  とすると、歪の期待値  $J$  は、全てのフレーム  $k$  に関して重み付け時の歪の中心を最小化するような (38) 式となる。

【0234】

【数42】

$$\begin{aligned}
 J &= \frac{1}{M} \sum_{k=1}^M \| \mathbf{W}_k' (\mathbf{y}_k - \mathbf{s}) \|^2 \\
 &= \frac{1}{M} \sum_{k=1}^M (\mathbf{y}_k - \mathbf{s})^T \mathbf{W}_k'^T \mathbf{W}_k' (\mathbf{y}_k - \mathbf{s}) \\
 &= \frac{1}{M} \sum_{k=1}^M \mathbf{y}_k^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{y}_k - 2 \mathbf{y}_k^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s} + \mathbf{s}^T \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s} \quad \dots (38)
 \end{aligned}$$

$$\frac{\partial J}{\partial \mathbf{s}} = \frac{1}{M} \sum_{k=1}^M (-2 \mathbf{y}_k^T \mathbf{W}_k'^T \mathbf{W}_k' + 2 \mathbf{s}^T \mathbf{W}_k'^T \mathbf{W}_k') = 0$$

を解いて

$$\sum_{k=1}^M \mathbf{y}_k^T \mathbf{W}_k'^T \mathbf{W}_k' = \sum_{k=1}^M \mathbf{s}^T \mathbf{W}_k'^T \mathbf{W}_k'$$

両辺の転置Tをとって

$$\sum_{k=1}^M \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{y}_k = \sum_{k=1}^M \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{s}$$

よって

$$\mathbf{s} = \left( \sum_{k=1}^M \mathbf{W}_k'^T \mathbf{W}_k' \right)^{-1} \sum_{k=1}^M \mathbf{W}_k'^T \mathbf{W}_k' \mathbf{y}_k \quad \dots (39)$$

【0235】上記(39)式で示す $\mathbf{s}$ は最適な代表ベクトルであり、最適なセントロイドコンディションである。

【0236】また、最適エンコード条件は、 $\| \mathbf{W}_k' (\mathbf{y}_k - \mathbf{s}) \|^2$ の値を最小化する $\mathbf{s}$ をサーチすればよい。ここでサーチ時の $\mathbf{W}_k'$ は、必ずしも学習時と同じ $\mathbf{W}_k'$ である必要はなく、重み無しで

【0237】

【数43】

$$\begin{bmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & & 1 \end{bmatrix}$$

【0238】のマトリクスとしてもよい。

【0239】このように、音声信号符号化装置内のベクトル量子化部116を2段のベクトル量子化部から構成することにより、出力するインデックスのビット数を可変にすることができる。

【0240】ところで、前述したように、スペクトルエンベロープ評価部148において得られるハーモニクススペクトルのデータ数は上記ピッチに応じて変化し、有効帯域が例えば3400kHzのときには8個から63個程度までのいずれかの個数のデータとなる。これらのデータをまとめてブロック化したベクトル $\mathbf{v}$ は可変次元ベクトルであり、上述の具体例では、ベクトル量子化の前に一定のデータ数、例えば44次元の固定次元の入力ベクトル $\mathbf{x}$ に次元変換している。この可変/固定次元変換は、上述したデータ数変換のことであり、具体的には例えば前述のように、オーバーサンプリング及び直線補間等を用いて実現することができる。

【0241】このような固定次元に変換したベクトル $\mathbf{x}$ に対して誤差計算を行って誤差を最小化するようなコードブックサーチを行うと、必ずしも元の可変次元ベクトル $\mathbf{v}$ に対する誤差を最小化するようなコードベクトルが選択されるとは限らない。

【0242】そこで、本実施の形態では、上記固定次元でのコードベクトルの選択を仮選択として複数のコードベクトルを選択するようにし、これらの仮選択された複数のコードベクトルについて、可変次元で最終的な最適コードベクトルの本選択を行わせるようにしている。なお、固定次元での仮選択を行わずに、可変次元での選択処理のみを行うようにしてもよい。

【0243】図10は、このような元の可変次元での最適ベクトル選択を行うための構成の一例を示しており、端子541には、上記スペクトルエンベロープ評価部148において得られるスペクトルエンベロープの可変個数のデータ、すなわち可変次元ベクトル $\mathbf{v}$ が入力されている。この可変次元の入力ベクトル $\mathbf{v}$ は、前述したデータ数変換回路である可変/固定次元変換回路542により、一定の個数、例えば44個のデータから成る固定次元(44次元)のベクトル $\mathbf{x}$ に変換され、端子501に送られている。この固定次元の入力ベクトル $\mathbf{x}$ と、固定次元の符号帳(コードブック)530から読み出される固定次元のコードベクトルとが固定次元の選択回路535に送られて、これらの間の重み付きの誤差あるいは歪が最小となるようなコードベクトルを符号帳530から選択するような選択処理あるいはコードブックサーチが行われる。

【0244】さらにこの図10の例においては、固定次元の符号帳530から得られた固定次元のコードベクトルを固定/可変次元変換回路544により元の可変次元

の入力ベクトル $\underline{v}$ と同じ可変次元に変換し、この可変次元に変換されたコードベクトルを可変次元の選択回路545に送って、上記入力ベクトル $\underline{v}$ との間の重み付き歪の計算を行い、その歪を最小とするコードベクトルを符号帳530から選択するような選択処理あるいはコードブックサーチを行っている。

【0245】すなわち、固定次元の選択回路535では、仮選択として、重み付き歪を最小化する候補となるいくつかのコードベクトルを選択しておき、これらの候補について、可変次元の選択回路545で重み付き歪計算を行って、歪を最小とするコードベクトルを本選択するようにしている。

【0246】この場合の仮選択及び本選択を用いるベクトル量子化についての適用範囲を簡単に説明する。このベクトル量子化は、ハーモニックコーディング、LPC残差のハーモニックコーディング、本件出願人が先に提案した特願平4-91422号明細書及び図面に開示したようなMBE（マルチバンド励起）符号化、LPC残差のMBE符号化等におけるハーモニクスベクトルに対して帯域制限型の次元変換を用いて可変次元のハーモニクスを重み付きベクトル量子化する場合に適用できるのみならず、その他、入力ベクトルの次元が可変であって、固定次元の符号帳を用いてベクトル量子化するようなあらゆる場合に適用できる。

【0247】上記仮選択としては、多段の量子化器構成の場合の一部を選択したり、シェイプコードブックとゲインコードブックとから成る符号帳の場合にシェイプコードブックのみを仮選択でサーチするようにしゲインについては可変次元での歪計算により決定するようにしたりすることが挙げられる。また、この仮選択について、前述した予備選択、すなわち、固定次元のベクトル $\underline{x}$ と符号帳に蓄えられた全てのコードベクトルとの類似度を近似計算（重み付き歪の近似計算）により求めて類似度の高い複数のコードベクトルを選択すること、を適用してもよい。この場合、固定次元での仮選択を上記予備選択とし、予備選択された候補のコードベクトルについて\*

$$E_1 = \| \underline{W}' (\underline{x} - g(\underline{s}_0 + \underline{s}_1)) \|^2 \quad \dots (b1)$$

で表される歪尺度 $E_1$ を最小化する $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ をサーチする。ここで、重み付け回路537での重み $\underline{W}'$  ※

$$\underline{W}' = \underline{W} \underline{H} / \| \underline{x} \| \quad \dots (b2)$$

であり、 $\underline{H}$ はLPC合成フィルタの周波数応答特性を対角要素に持つマトリクスを、また $\underline{W}$ は聴覚重み付けフィルタの周波数応答特性を対角要素に持つマトリクスをそれぞれ示している。

【0252】先ず、上記(b1)式の歪尺度 $E_1$ を最小★

$$E_2 = \| \underline{W}_v (\underline{v} - D_2 g(\underline{s}_0 + \underline{s}_1)) \|^2 \quad \dots (b3)$$

を最小化する $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ の組を最適コードベクトルとして、最終的な本選択を可変次元で行う。

【0253】上記(b1)式についてのサーチ、学習については、前述した(27)式以下の説明の通りであ

\*可変次元での重み付き歪を最小化するような本選択を行わせてもよく、また、仮選択の工程で上記予備選択のみならず高精度の歪演算による絞り込みをさらに行った後に本選択に回すようにしてもよい。

【0248】以下、このような仮選択及び本選択を用いたベクトル量子化の具体例について、図面を参照しながら説明する。

【0249】図10においては、符号帳530は、シェイプコードブック531とゲインコードブック532とから成り、シェイプコードブック531は、さらに2つのコードブックCB0、CB1を有している。これらのシェイプコードブックCB0、CB1からの出力コードベクトルをそれぞれ $\underline{s}_0$ 、 $\underline{s}_1$ とし、ゲインコードブック532により決定されるゲイン回路533のゲインを $g$ とする。入力端子541からの可変次元の入力ベクトル $\underline{v}$ は、可変／固定次元変換回路542により次元変換（これを $D_1$ とする）されて、端子501を介して固定次元のベクトル $\underline{x}$ として選択回路535の減算器536に送られ、符号帳530から読み出された固定次元のコードベクトルとの差がとられ、重み付け回路537により重み付けがなされて、誤差最小化回路538に送られる。この重み付け回路537での重みを $\underline{W}'$ とする。また、符号帳530から読み出された固定次元のコードベクトルは、固定／可変次元変換回路544により次元変換（これを $D_2$ とする）されて、可変次元の選択回路545の減算器546に送られ、可変次元の入力ベクトル $\underline{v}$ との差がとられ、重み付け回路547により重み付けがなされて、誤差最小化回路548に送られる。この重み付け回路547での重みを $\underline{W}_v$ とする。

【0250】ここで、誤差最小化回路538、548の誤差とは、上記歪あるいは歪尺度のことであり、誤差すなわち歪が小さくなることは、類似度あるいは相関性が高まることに相当する。

【0251】固定次元での上記仮選択を行う選択回路535では、前記(27)式の説明と同様に、

★化する $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ をサーチする。ここで、 $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ を、上記歪尺度 $E_1$ を小さくする順に、上位からL組とっておき（固定次元での仮選択）、そのL組の $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ に関して、

【0254】以下、上記(b3)式に基づくコードブック学習のためのセントロイドコンディションについて説明する。



【0255】符号帳(コードブック)530の内のシェイプコードブック531の1つであるコードブックCB0について、コードベクトル $\underline{s}_0$ を選択する全てのフレーム $k$ に関して、歪の期待値を最小化する。そのようなフレームが $M$ 個あるとして、

【0256】

【数44】

$$\frac{\partial J}{\partial \underline{s}_{0c}} = \frac{1}{M} \sum_{k=1}^M \left\{ -2g_k D_{2k}^T W_{vk}^T W_{vk} \underline{v}_k + 2g_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0c} + 2g_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{1k} \right\} = 0 \quad \cdots (b4)$$

【0259】を解いて、

【0260】

※【数46】

※

$$\underline{s}_{0c} = \left\{ \sum_{k=1}^M g_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \right\}^{-1} \times \sum_{k=1}^M \left\{ g_k D_{2k}^T W_{vk}^T W_{vk} (\underline{v}_k - g_k D_{2k} \underline{s}_{1k}) \right\} \quad \cdots (b6)$$

【0261】となる。この(b6)式で、 $\{ \}^{-1}$ は逆行列を、 $W_{vk}^T$ は $W_{vk}$ の転置行列をそれぞれ示している。この(b6)式が、シェイプベクトル $\underline{s}_0$ の最適セントロイド条件である。

【0262】次に、符号帳(コードブック)530の内のシェイプコードブック531のもう1つのコードブックCB1についてのコードベクトル $\underline{s}_1$ を選択する場合も同様であるため、説明を省略する。

【0263】次に、符号帳(コードブック)530の内のゲインコードブック532からのゲイン $g$ についてのセントロイド条件を考察する。

★30

$$\frac{\partial J_g}{\partial g_c} = \frac{1}{M} \sum_{k=1}^M \left\{ -2\underline{v}_{vk}^T W_{vk}^T W_{vk} D_{2k} (\underline{s}_{0k} + \underline{s}_{1k}) + 2g_c (D_{2k} (\underline{s}_{0k} + \underline{s}_{1k}))^T W_{vk}^T W_{vk} D_{2k} (\underline{s}_{0k} + \underline{s}_{1k}) \right\} = 0 \quad \cdots (b8)$$

【0268】を解いて、

【0269】

☆【数49】

☆

$$g_c = \frac{\sum_{k=1}^M \underline{v}_{vk}^T W_{vk}^T W_{vk} D_{2k} (\underline{s}_{0k} + \underline{s}_{1k})}{\sum_{k=1}^M (D_{2k} (\underline{s}_{0k} + \underline{s}_{1k}))^T W_{vk}^T W_{vk} D_{2k} (\underline{s}_{0k} + \underline{s}_{1k})} \quad \cdots (b9)$$

【0270】これがゲインのセントロイド条件である。

【0271】次に、上記(b3)式に基づく最適エンコード条件を考察する。上記(b3)式でサーチせねばならない $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ の組は、上記固定次元での仮選択により $L$ 組と限定されているので、上記(b3)式を $L$ 組の $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ に関して直接計算し、歪 $E_2$ を最小とする $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ の組を最適コードベクトルとして選択す◆

◆ればよい。

【0272】ここで、仮選択の $L$ が非常に大きい場合や、上記仮選択を行わず直接的に可変次元で $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ の選択を行う場合に、有効とされるシェイプとゲインのシーケンシャルなサーチの方法について説明する。

【0273】上記(b3)式の各 $\underline{s}_0$ 、 $\underline{s}_1$ 、 $g$ に、それぞれ $i$ 、 $j$ 、 $l$ のインデックスを付加して書き直すと、

$$E_2 = \| W_v (\underline{v} - D_2 g_l (\underline{s}_{0i} + \underline{s}_{1j})) \|^2 \quad \cdots (b10)$$

$$J = \frac{1}{M} \sum_{k=1}^M \| W_{vk} (\underline{v}_k - g_k D_{2k} (\underline{s}_{0c} + \underline{s}_{1k})) \|^2 \quad \cdots (b4)$$

【0257】を最小化すればよい。この(b4)式を最小化するために、

【0258】

【数45】

となる。これを最小化する  $g_1, \underline{s}_{0i}, \underline{s}_{1j}$  を総当たりでサーチすることも可能であるが、例えば  $0 \leq i < 3$

$2, 0 \leq j < 3, 2, 0 \leq i < 3, 2$  とすると、 $3 \times 2 = 6$  のパターンについて上記 (b10) 式を計算することになり、膨大な演算量となる。そこで、シエ \*

$$E_1 = \|\mathbf{W}_V(\underline{v} - D_1 g_1 \underline{s}_*)\|^2 \quad \dots (b11)$$

と表せ、さらに、 $\underline{v}_w = \mathbf{W}_V \underline{v}$ ,  $\underline{s}_w = \mathbf{W}_V D_1 \underline{s}_*$  とお

くと、(b11) 式は、

$$\begin{aligned} E_2 &= \|\underline{v}_w - g_1 \underline{s}_w\|^2 \\ &= \|\underline{v}_w\|^2 + \|\underline{s}_w\|^2 \left( g_1 - \frac{\underline{v}_w^T \cdot \underline{s}_w}{\|\underline{s}_w\|^2} \right)^2 - \left( \frac{\underline{v}_w^T \cdot \underline{s}_w}{\|\underline{s}_w\|^2} \right)^2 \quad \dots (b1) \end{aligned}$$

【0276】となる。よって、 $g_1$  の精度が充分にとれ

るとすると、

【0277】

【数51】

$$\frac{(\underline{v}_w^T \cdot \underline{s}_w)^2}{\|\underline{s}_w\|^2} \quad \dots (b13)$$

を最大化する  $\underline{s}_w$  をサーチ

$$\frac{\underline{v}_w^T \cdot \underline{s}_w}{\|\underline{s}_w\|^2} \quad \dots (b14)$$

にもっとも近い  $g_1$  をサーチ

【0278】元の変数を代入して書き直すと、次の (b15)、(b16) 式ようになる。

【0279】

【数52】

$$\frac{(\underline{v}^T \mathbf{W}_V^T \mathbf{W}_V D_2 (\underline{s}_{0i} + \underline{s}_{1j}))^2}{\|\mathbf{W}_V D_2 (\underline{s}_{0i} + \underline{s}_{1j})\|^2} \quad \dots (b15)$$

を最大化する  $\underline{s}_{0i}, \underline{s}_{1j}$  の組をサーチ

$$\frac{\underline{v}^T \mathbf{W}_V^T \mathbf{W}_V D_2 (\underline{s}_{0i} + \underline{s}_{1j})}{\|\mathbf{W}_V D_2 (\underline{s}_{0i} + \underline{s}_{1j})\|^2} \quad \dots (b16)$$

にもっとも近い  $g_1$  をサーチ

【0280】上記 (b6)、(b9) 式のシェイプ、ゲインのセントロイド条件と、上記 (b15)、(b16) 式の最適エンコード条件 (Nearest Neighbour Condition) を用いて、一般化ロイドアルゴリズム (Generalized Lloyd Algorithm: GLA) によって、コードブック (CB0、CB1、CBg) を同時に学習させることができる。

【0281】これらの (b6)、(b9)、(b15)、(b16) 式を用いた学習法は、先に述べた (27) 式以下の説明、特に前記 (31)、(32)、(35) 式を用いる方法に比べて、元の入力ベクトル  $\underline{v}$  の可変次元への変換を行った後の歪を最小化している点で優れてい

\* イブとゲインをシーケンシャルにサーチする方法を説明する。

【0274】まず、シェイプコードベクトル  $\underline{s}_{0i}, \underline{s}_{1j}$  を決定してから、ゲイン  $g_1$  を決定する。 $\underline{s}_{0i} + \underline{s}_{1j} = \underline{s}_*$  とおくと、上記 (b10) 式は、

※【0275】

※【数50】

る。

【0282】しかし、上記 (b6)、(b9) 式、特に (b6) 式の演算は、煩雑であるので、例えば上記 (b15)、(b16) 式の最適エンコード条件のみを用いて、セントロイド条件は前記 (27) 式 (すなわち (b1) 式) の最適化から導かれるものを用いてもよい。

20 【0283】あるいは、コードブックの学習時は、全て前記 (27) 式以下の説明に述べた方法で行い、サーチ時のみ上記 (b15)、(b16) 式を用いる方法も挙げられる。また、上記固定次元での仮選択を前記 (27) 式以下の説明に述べた方法で行い、選ばれた複数個 (L 個) の組についてのみ上記 (b3) 式を直接評価してサーチを行うようにしてもよい。

【0284】いずれにしても、上記 (b3) 式の歪評価によるサーチを、上記仮選択後、あるいは総当たりに使用することにより、最終的にはより歪の少ないコードベクトルサーチあるいは学習を行うことが可能となる。

30

【0285】ここで、元の入力ベクトル  $\underline{v}$  と同じ可変次元で歪計算を行うことが好ましい理由について簡単に述べる。

【0286】これは、固定次元での歪の最小化と可変次元での歪の最小化とが一致すれば、可変次元での歪の最小化は不要であるが、固定/可変次元変換回路544での次元変換D<sub>1</sub>が直交行列ではないため、これらの歪の最小化は一致しない。このため、固定次元で歪を最小化しても、必ずしもこれは可変次元で最適に歪を最小化することにはならず、最終的に得られる可変次元のベクトルを最適化しようとするには、可変次元での最適化が必要とされるからである。

【0287】次に図11は、符号帳(コードブック)をシェイプコードブックとゲインコードブックとに分けるときのゲインを可変次元でのゲインとし、可変次元で最適化するようにした例を示している。

【0288】すなわち、シェイプコードブック531から読み出された固定次元のコードベクトルを固定/可変次元変換回路544に送って可変次元のベクトルに変換した後、ゲイン回路533に送っている。可変次元での

50

選択回路545は、ゲイン回路533からの可変次元のコードベクトルと上記入力ベクトル $\underline{v}$ に基づいて、固定/可変次元変換されたコードベクトルに対するゲイン回路533での最適ゲインを選択すればよい。あるいは、ゲイン回路533への入力ベクトルと上記入力ベクトル $\underline{v}$ との内積に基づいて最適ゲインを選択するようにしてもよい。他の構成及び動作は、上記図10の例と同様である。

【0289】なお、シェイプコードブック531については、選択回路535における固定次元での選択時に唯一のコードベクトルを選択するようにし、可変次元での選択はゲインのみとしてもよい。

【0290】このように、固定/可変次元変換回路544で変換したコードベクトルに対してゲインを掛けるような構成とすることにより、上記図10に示すようなゲイン倍したコードベクトルを固定/可変次元変換するものに比べて、固定/可変次元変換による影響を考慮した上で最適なゲインを選択することができる。

【0291】次に、このような固定次元での仮選択と可変次元での本選択とを組み合わせるベクトル量子化の他の具体例について説明する。

【0292】以下の具体例では、第1の符号帳から読み出された固定次元の第1のコードベクトルを入力ベクトルの可変次元に次元変換し、第2の符号帳から読み出された固定次元の第2のコードベクトルを上記固定/可変次元変換された可変次元の第1のコードベクトルに加算し、この加算されて得られた加算コードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを上記少なくとも第2の符号帳より選択するよ\*

$$E_3 = \| \mathbf{W}_v (\underline{v} - (D_2 g \underline{s}_0 + \underline{s}_1)) \|^2 \cdots (b17)$$

となる。

【0296】次に、図13の例では、ゲイン回路533を加算器549の出力側に配置している。従って、第1の符号帳CB0から読み出され固定/可変次元変換回路544で可変次元に変換されたコードベクトルと、第2の符号帳CB1から読み出されたコードベクトルとの加\*

$$E_4 = \| \mathbf{W}_v (\underline{v} - g (D_2 \underline{s}_0 + \underline{s}_1)) \|^2 \cdots (b18)$$

となる。この図13の例の他の構成は、上記図12の例と同様であるため説明を省略する。

【0297】次に、図14の例では、上記図12の例における第1の符号帳CB0の出力側にゲイン $g$ のゲイン回路533を設けるのみならず、第2の符号帳CB1の出力側にもゲイン $g$ のゲイン回路533を設けている。この図14の選択回路545で計算される歪は、上記図13の例と同様に、(b18)式に示す歪 $E_4$ となる。この図14の例の他の構成は、上記図12の例と同様であるため説明を省略する。

$$E_5 = \| \mathbf{W}_v (\underline{v} - (g D_2 (\underline{s}_0 + \underline{s}_1) + \underline{s}_2)) \|^2 \cdots (b19)$$

となる。この図15の例の他の構成は、上記図12の例と同様であるため説明を省略する。

\*うにしている。

【0293】例えば、図12の例では、第1の符号帳(コードブック)CB0から読み出された固定次元の第1のコードベクトル $\underline{s}_0$ を、固定/可変次元変換回路544に送って、端子541の入力ベクトル $\underline{v}$ に等しい可変次元に次元変換し、第2の符号帳CB1から読み出された固定次元の第2のコードベクトル $\underline{s}_1$ を加算器549に送って、固定/可変次元変換回路544からの可変次元のコードベクトルに加算し、この加算器549で加算されて得られた加算コードベクトルを選択回路545に送り、この選択回路545で、加算器549からの加算ベクトルと上記入力ベクトル $\underline{v}$ との誤差を最小化する最適のコードベクトルを選択するようにしている。ここで、第2の符号帳(コードブック)CB1からのコードベクトルは、入力ベクトルのハーモニクスの低域側からコードブックCB1の次元にまで適用される。また、ゲイン $g$ のゲイン回路533は、第1の符号帳CB0と固定/可変次元変換回路544との間にのみ設けられている。他の構成は上記図10と同様であるため、対応する部分に同じ指示符号を付して説明を省略する。

【0294】このように、コードブックCB1からの固定次元のままのコードベクトルを、コードブックCB0から読み出されて可変次元に変換されたコードベクトルと加算することにより、固定/可変次元変換を行うことによって発生した歪をコードブックCB1からの固定次元のコードベクトルによって減じることができる。

【0295】この図12の可変次元の選択回路545で計算される歪 $E_3$ は、

※算結果に対してゲイン $g$ が掛けられる。これは、CB0からのコードベクトルに乗ずるべきゲインと、その補正分(量子化誤差の量子化)のためのコードブックCB1からのコードベクトルに乗ずるべきゲインの相関が強いいため、共通のゲインを用いている。この図13の選択回路545で計算される歪 $E_4$ は、

★【0298】次に、図15は、上記図12の第1の符号帳を2つのシェイプコードブックCB0、CB1で構成した例を示し、これらのシェイプコードブックCB0、CB1からの各コードベクトル $\underline{s}_0$ 、 $\underline{s}_1$ が加算され、ゲイン回路533でゲイン $g$ を掛けられて、固定/可変次元変換回路544に送られている。この固定/可変次元変換回路544からの可変次元のコードベクトルと、第2の符号帳CB2からのコードベクトル $\underline{s}_2$ を加算器549で加算して、選択回路545に送っている。この図15の選択回路545で計算される歪 $E_5$ は、

【0299】ここで、上記(b18)式におけるサーチ方法について説明する。先ず、第1のサーチ方法として

は、

$$E_4' = \| \mathbf{W}' (\mathbf{x} - g_1 \mathbf{s}_{01}) \|^2 \quad \dots (b20)$$

を最小化する  $\mathbf{s}_{01}$ ,  $g_1$  をサーチし、次に

$$E_4 = \| \mathbf{W}_v (\mathbf{y} - g_1 (D_2 \mathbf{s}_{01} + \mathbf{s}_{11})) \|^2 \quad \dots (b21)$$

を最小化する  $\mathbf{s}_{11}$  をサーチすることが挙げられる。

【0309】

【0300】第2のサーチ方法としては、

【数56】

【0301】

【数53】

$$\frac{(\mathbf{s}_{01}^T \mathbf{W}'^T \mathbf{W}' \mathbf{x})^2}{\|\mathbf{W}' \mathbf{s}_{01}\|^2} \quad \dots (b22)$$

を最大化する  $\mathbf{s}_{01}$  をサーチ

$$\frac{(\mathbf{y}^T \mathbf{W}_v^T \mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11}))^2}{\|\mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11})\|^2} \quad \dots (b23)$$

を最大化する  $\mathbf{s}_{11}$  をサーチ

$$\frac{\mathbf{y}^T \mathbf{W}_v^T \mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11})}{\|\mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11})\|^2} \quad \dots (b24)$$

にもっとも近いゲイン  $g_1$  をサーチ

【0302】が挙げられる。

【0303】第3のサーチ方法としては、

【0304】

【数54】

$$E_4' = \|\mathbf{W}' (\mathbf{x} - g_1 \mathbf{s}_{01})\|^2 \quad \dots (b25)$$

を最小化する  $\mathbf{s}_{01}$ ,  $g_1$  をサーチ

$$\frac{(\mathbf{y}^T \mathbf{W}_v^T \mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11}))^2}{\|\mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11})\|^2} \quad \dots (b26)$$

を最大化する  $\mathbf{s}_{11}$  をサーチ

$$\frac{\mathbf{y}^T \mathbf{W}_v^T \mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11})}{\|\mathbf{W}_v (D_2 \mathbf{s}_{01} + \mathbf{s}_{11})\|^2} \quad \dots (b27)$$

にもっとも近いゲイン  $g_1$  を本選択

【0305】が挙げられる。

【0306】次に、上記第1のサーチ方法の上記 (b20) 式のセントロイド条件について説明する。上記コードベクトル  $\mathbf{s}_{01}$  のセントロイドを  $\mathbf{s}_{0c}$  とするとき、

【0307】

【数55】

$$J = \frac{1}{M} \sum_{k=1}^M \|\mathbf{W}'_k (\mathbf{x}_k - g_k \mathbf{s}_{0c})\|^2 \quad \dots (b28)$$

【0308】を最小化する。これを最小化するために、

$$\frac{\partial J}{\partial \mathbf{s}_{0c}} = \frac{1}{M} \sum_{k=1}^M (-2g_k \mathbf{W}'_k^T \mathbf{W}'_k \mathbf{x}_k + 2g_k^2 \mathbf{W}'_k^T \mathbf{W}'_k \mathbf{s}_{0c}) = 0 \quad \dots (b29)$$

【0310】を解いて、

【0311】

【数57】

$$\mathbf{s}_{0c} = \left\{ \sum_{k=1}^M g_k^2 \mathbf{W}'_k^T \mathbf{W}'_k \right\}^{-1} \sum_{k=1}^M g_k \mathbf{W}'_k^T \mathbf{W}'_k \mathbf{x}_k \quad \dots (b30)$$

【0312】が得られる。同様に、ゲイン  $g$  のセントロイド  $g_c$  については、上記 (b20) 式より、

【0313】

20 【数58】

$$J = \frac{1}{M} \sum_{k=1}^M \|\mathbf{W}'_k (\mathbf{x}_k - g_c \mathbf{s}_{0c})\|^2 \quad \dots (b31)$$

【0314】

【数59】

$$\frac{\partial J}{\partial g_c} = \frac{1}{M} \sum_{k=1}^M (-2\mathbf{s}_{0c}^T \mathbf{W}'_k^T \mathbf{W}'_k \mathbf{x}_k + 2g_c \mathbf{s}_{0c}^T \mathbf{W}'_k^T \mathbf{W}'_k \mathbf{s}_{0c}) = 0 \quad \dots (b32)$$

30 【0315】を解いて、

【0316】

【数60】

$$g_c = \frac{\sum_{k=1}^M \mathbf{s}_{0c}^T \mathbf{W}'_k^T \mathbf{W}'_k \mathbf{x}_k}{\sum_{k=1}^M \mathbf{s}_{0c}^T \mathbf{W}'_k^T \mathbf{W}'_k \mathbf{s}_{0c}} \quad \dots (b33)$$

【0317】また、上記第1のサーチ方法の上記 (b21) 式のセントロイド条件として、ベクトル  $\mathbf{s}_{11}$  のセントロイド  $\mathbf{s}_{1c}$  については、

【0318】

【数61】

$$J = \frac{1}{M} \sum_{k=1}^M \|\mathbf{W}_{vk} (\mathbf{y}_k - g_k (D_2 \mathbf{s}_{0c} + \mathbf{s}_{1c}))\|^2 \quad \dots (b34)$$

【0319】

【数62】

$$\frac{\partial J}{\partial \underline{s}_{1c}} = \frac{1}{M} \sum_{k=1}^M \{ -2g_k \mathbf{W}_{vk}^T \mathbf{W}_{vk} \underline{v}_k + 2g_k^2 \mathbf{W}_{vk}^T \mathbf{W}_{vk} D_{2k} \underline{s}_{0k} + 2g_k^2 \mathbf{W}_{vk}^T \mathbf{W}_{vk} \underline{s}_{1k} \} = 0 \quad \cdots (b35)$$

【0320】を解いて、

＊【数63】

【0321】

＊

$$\underline{s}_{1c} = \left\{ \sum_{k=1}^M g_k^2 \mathbf{W}_{vk}^T \mathbf{W}_{vk} \right\}^{-1} \sum_{k=1}^M g_k \mathbf{W}_{vk}^T \mathbf{W}_{vk} (\underline{v}_k - g_k D_{2k} \underline{s}_{0k}) \quad \cdots (b36)$$

【0322】が得られる。上記(b21)式から上記コー 10＊【0324】

ドベクトル  $\underline{s}_{01}$  のセントロイド  $\underline{s}_{0c}$  を求めると、

【数65】

【0323】

【数64】

$$J = \frac{1}{M} \sum_{k=1}^M \|\mathbf{W}_{vk} (\underline{v}_k - g_k (D_{2k} \underline{s}_{0c} + \underline{s}_{1k}))\|^2 \quad \cdots (b37)$$

※

$$\frac{\partial J}{\partial \underline{s}_{0c}} = \frac{1}{M} \sum_{k=1}^M \{ -2g_k D_{2k}^T \mathbf{W}_{vk}^T \mathbf{W}_{vk} \underline{v}_k + 2g_k^2 D_{2k}^T \mathbf{W}_{vk}^T \mathbf{W}_{vk} D_{2k} \underline{s}_{0c} + 2g_k^2 D_{2k}^T \mathbf{W}_{vk}^T \mathbf{W}_{vk} \underline{s}_{1k} \} = 0 \quad \cdots (b38)$$

【0325】

★ ★【数66】

$$\underline{s}_{0c} = \left\{ \sum_{k=1}^M g_k^2 D_{2k}^T \mathbf{W}_{vk}^T \mathbf{W}_{vk} D_{2k} \right\}^{-1} \sum_{k=1}^M g_k D_{2k}^T \mathbf{W}_{vk}^T \mathbf{W}_{vk} (\underline{v}_k - g_k \underline{s}_{1k}) \quad \cdots (b39)$$

【0326】が得られる。同様に、上記(b21)式から ☆【0327】

上記ゲイン  $g$  のセントロイド  $g_c$  を求めると、

☆ 【数67】

$$g_c = \frac{\sum_{k=1}^M \underline{v}_k^T \mathbf{W}_{vk}^T \mathbf{W}_{vk} (D_{2k} \underline{s}_{0k} + \underline{s}_{1k})}{\sum_{k=1}^M (D_{2k} \underline{s}_{0k} + \underline{s}_{1k})^T \mathbf{W}_{vk}^T \mathbf{W}_{vk} (D_{2k} \underline{s}_{0k} + \underline{s}_{1k})} \quad \cdots (b40)$$

【0328】が得られる。

【0329】以上、上記(b20)式によるコードベクトル  $\underline{s}_{01}$  のセントロイド  $\underline{s}_{0c}$  の算出方法を(b30)式に、ゲイン  $g$  のセントロイド  $g_c$  の算出方法を(b33)式にそれぞれ示した。また、上記(b21)式によるセントロイドの算出方法として、コードベクトル  $\underline{s}_{11}$  のセントロイド  $\underline{s}_{1c}$  を(b36)式に、コードベクトル  $\underline{s}_{01}$  のセントロイド  $\underline{s}_{0c}$  を(b39)式に、ゲイン  $g$  のセントロイド  $g_c$  を(b40)式にそれぞれ示した。

【0330】実際の一般化ロイドアルゴリズム(GL

A)によるコードブックの学習においては、セントロイ◆

◆ド条件として、上記(b30)式、(b36)式、(b40)式を使用して  $\underline{s}_{01}$ 、 $\underline{s}_{11}$ 、 $g$  を同時に学習する方法が挙げられる。サーチ方法(Nearest Neighbour Condition)は、例えば上記(b22)式、(b23)式、(b24)式を用いればよい。その他、上記(b30)式、(b33)式、(b36)式、あるいは、上記(b39)式、(b36)式、(b40)式といったセントロイド条件の組み合わせも可能であることは勿論である。

【0331】次に、上記図12に対応する上記(b17)式の歪尺度の場合のサーチ方法について説明する。この場合には、

$$E_3' = \|\mathbf{W}' (\underline{x} - g_1 \underline{s}_{01})\|^2 \quad \cdots (b41)$$

を最小化する  $\underline{s}_{01}$ 、 $g_1$  をサーチし、次に

$$E_3 = \|\mathbf{W}_v (\underline{v} - g_1 (D_2 \underline{s}_{01} + \underline{s}_{11}))\|^2 \quad \cdots (b42)$$

を最小化する  $\underline{s}_{11}$  をサーチすることが挙げられる。

にしている。

【0332】上記(b41)式において、全ての  $g_1$ 、 $\underline{s}_{01}$  の組を総当たりするのは現実的でないので、次のよう

【0333】

【数68】

$$\frac{(\underline{x}^T \mathbf{W}'^T \mathbf{W}' \underline{s}_{0i})^2}{\|\mathbf{W}' \underline{s}_{0i}\|^2} \dots (b43)$$

を最大化する  $\underline{s}_{0i}$  を上位  $L$  個サーチ

$$\frac{\underline{x}^T \mathbf{W}'^T \mathbf{W}' \underline{s}_{0i}}{\|\mathbf{W}' \underline{s}_{0i}\|^2} \dots (b44)$$

にもっとも近いゲインを上記 (b43) 式に対応して  $L$  個サーチ

$$E_3 = \|\mathbf{W}_V (\underline{y} - (D_2 \underline{s}_{0i} + \underline{s}_{1j}))\|^2 \dots (b45)$$

を最小化する  $\underline{s}_{1j}$  をサーチ

【0334】次に、上記 (b41) 式、(b42) 式よりセントロイド条件を導く。この場合も、上述したのと同様に、どの式を用いるかで変わってくる。

【0335】まず、上記 (b41) 式を用いる場合には、上記コードベクトル  $\underline{s}_{01}$  のセントロイドを  $\underline{s}_{0c}$  とすると

き、

【0336】

【数69】

$$J = \frac{1}{M} \sum_{k=1}^M \|\mathbf{W}'_k (\underline{x}_k - g_k \underline{s}_{0c})\|^2 \dots (b46)$$

【0337】を最小化することにより、

【0338】

【数70】

$$\underline{s}_{0c} = \left\{ \sum_{k=1}^M g_k^2 \mathbf{W}'_k^T \mathbf{W}'_k \right\}^{-1} \sum_{k=1}^M g_k \mathbf{W}'_k^T \mathbf{W}'_k \underline{x}_k \dots (b47)$$

【0339】が得られる。同様に、ゲイン  $g$  のセントロイド  $g_c$  については、上記 (b41) 式より、上記 (b3 \*

$$\frac{\partial J}{\partial \underline{s}_{1c}} = \frac{1}{M} \sum_{k=1}^M \left\{ 2 \mathbf{W}_{V_k}^T \mathbf{W}_{V_k} \underline{s}_{1c} - 2 \mathbf{W}_{V_k}^T \mathbf{W}_{V_k} \underline{y}_k + 2 g_k \mathbf{W}_{V_k}^T \mathbf{W}_{V_k} D_{2k} \underline{s}_{0k} \right\} = 0 \dots (b50)$$

【0344】を解いて、

【0345】

※【数74】

※

$$\underline{s}_{1c} = \left\{ \sum_{k=1}^M \mathbf{W}_{V_k}^T \mathbf{W}_{V_k} \right\}^{-1} \sum_{k=1}^M \mathbf{W}_{V_k}^T \mathbf{W}_{V_k} (\underline{y}_k - g_k D_{2k} \underline{s}_{0k}) \dots (b51)$$

【0346】が得られる。同様に、上記 (b42) 式から

上記コードベクトル  $\underline{s}_{01}$  のセントロイド  $\underline{s}_{0c}$ 、及び上記

ゲイン  $g$  のセントロイド  $g_c$  を求めることができる。★

$$\underline{s}_{0c} = \left\{ \sum_{k=1}^M g_k^2 D_{2k}^T \mathbf{W}_{V_k}^T \mathbf{W}_{V_k} D_{2k} \right\}^{-1} \sum_{k=1}^M g_k D_{2k}^T \mathbf{W}_{V_k}^T \mathbf{W}_{V_k} (\underline{y}_k - \underline{s}_{1k}) \dots (b52)$$

【0348】

【数76】

\*3) 式の場合と同様に、次の式が得られる。

【0340】

【数71】

$$g_c = \frac{\sum_{k=1}^M \underline{s}_{0k}^T \mathbf{W}_k^T \mathbf{W}_k \underline{x}_k}{\sum_{k=1}^M \underline{s}_{0k}^T \mathbf{W}_k^T \mathbf{W}_k \underline{s}_{0k}} \dots (b48)$$

20

【0341】また、上記 (b42) 式を用いてベクトル  $\underline{s}_{11}$  のセントロイド  $\underline{s}_{1c}$  を求める場合には、次の通りである。

【0342】

【数72】

$$J = \frac{1}{M} \sum_{k=1}^M \|\mathbf{W}_{V_k} (\underline{y}_k - D_{2k} g_k \underline{s}_{0k} - \underline{s}_{1c})\|^2 \dots (b49)$$

【0343】

【数73】

30

★【0347】

【数75】

$$J = \frac{1}{M} \sum_{k=1}^M \left\{ y_k^T W_{vk}^T W_{vk} y_k + g_c^2 (D_{2k} \varepsilon_{0k})^T W_{vk}^T W_{vk} D_{2k} \varepsilon_{0k} \right. \\ \left. + \varepsilon_{1k}^T W_{vk}^T W_{vk} \varepsilon_{1k} - 2g_c (D_{2k} \varepsilon_{0k})^T W_{vk}^T W_{vk} y_k \right. \\ \left. - 2\varepsilon_{1k}^T W_{vk}^T W_{vk} y_k + 2\varepsilon_{1k}^T W_{vk}^T W_{vk} D_{2k} \varepsilon_{0k} g_c \right\} \cdots (b53)$$

【0349】

\* \* 【数77】

$$\frac{\partial J}{\partial g_c} = \frac{1}{M} \sum_{k=1}^M \left\{ 2g_c (D_{2k} \varepsilon_{0k})^T W_{vk}^T W_{vk} D_{2k} \varepsilon_{0k} \right. \\ \left. - 2(D_{2k} \varepsilon_{0k})^T W_{vk}^T W_{vk} y_k + 2\varepsilon_{1k}^T W_{vk}^T W_{vk} D_{2k} \varepsilon_{0k} \right\} \\ = 0 \cdots (b54)$$

【0350】

【数78】

$$g_c = \frac{\sum_{k=1}^M \left\{ (D_{2k} \varepsilon_{0k})^T W_{vk}^T W_{vk} (y_k - \varepsilon_{1k}) \right\}}{\sum_{k=1}^M (D_{2k} \varepsilon_{0k})^T W_{vk}^T W_{vk} D_{2k} \varepsilon_{0k}} \cdots (b55)$$

【0351】なお、一般化ロイドアルゴリズム (GLA) によるコードブックの学習は、上記 (b47) 式、(b48) 式、(b51) 式を用いて、あるいは、上記 (b51) 式、(b52) 式、(b55) 式を用いて行うようにすればよい。

【0352】次に、本発明の前記 CELP 符号化構成を用いた第2の符号化部120は、より具体的には図16に示すような、多段のベクトル量子化処理部 (図16の例では2段の符号化部120<sub>1</sub>と120<sub>2</sub>) の構成を有するものとなされている。なお、当該図16の構成は、伝送ビットレートを例えば前記2kbpsと6kbpsとで切り換え可能な場合において、6kbpsの伝送ビットレートに対応した構成を示しており、さらにシェイプ及びゲインインデックス出力を23ビット/5msecと15ビット/5msecとで切り換えられるようにしているものである。また、この図16の構成における処理の流れは図17に示すようになっている。

【0353】この図16において、例えば、図16の第1の符号化部300は前記図3の第1の符号化部113と略々対応し、図16のLPC分析回路302は前記図3に示したLPC分析回路132と対応し、図16のLSPパラメータ量子化回路303は図3の前記 $\alpha$ →LSP変換回路133からLSP→ $\alpha$ 変換回路137までの構成と対応し、図16の聴覚重み付けフィルタ304は図3の前記聴覚重み付けフィルタ算出回路139及び聴覚重み付けフィルタ125と対応している。したがって、この図16において、端子305には前記図3の第1の符号化部113のLSP→ $\alpha$ 変換回路137からの出力と同じものが供給され、また、端子307には前記図3の聴覚重み付けフィルタ算出回路139からの出力と同じものが、端子306には前記図3の聴覚重み付け

フィルタ125からの出力と同じものが供給される。ただし、この図16の聴覚重み付けフィルタ304では、前記図3の聴覚重み付けフィルタ125とは異なり、前記LSP→ $\alpha$ 変換回路137の出力を用いずに、入力音声データと量子化前の $\alpha$ パラメータとから、前記聴覚重み付けした信号 (すなわち前記図3の聴覚重み付けフィルタ125からの出力と同じ信号) を生成している。

【0354】また、この図16に示す2段構成の第2の符号化部120<sub>1</sub>及び120<sub>2</sub>において、減算器313及び323は図3の減算器123と対応し、距離計算回路314及び324は図3の距離計算回路124と、ゲイン回路311及び321は図3のゲイン回路126と、ストキャスティックコードブック310、320及びゲインコードブック315、325は図3の雑音符号帳121とそれぞれ対応している。

【0355】このような図16の構成において、まず、図17のステップS1に示すように、LPC分析回路302では、端子301から供給された入力音声データ $x$ を前述同様に適当なフレームに分割してLPC分析を行い、 $\alpha$ パラメータを求める。LSPパラメータ量子化回路303では、上記LPC分析回路302からの $\alpha$ パラメータをLSPパラメータに変換して量子化し、さらにこの量子化したLSPパラメータを補間した後、 $\alpha$ パラメータに変換する。次に、当該LSPパラメータ量子化回路303では、当該量子化したLSPパラメータを変換した $\alpha$ パラメータ、すなわち量子化された $\alpha$ パラメータから、LPC合成フィルタ関数 $1/H(z)$ を生成し、これを端子305を介して1段目の第2の符号化部120<sub>1</sub>の聴覚重み付け合成フィルタ312に送る。

【0356】一方、聴覚重み付けフィルタ304では、LPC分析回路302からの $\alpha$ パラメータ (すなわち量子化前の $\alpha$ パラメータ) から、前記図3の聴覚重み付けフィルタ算出回路139によるものと同じ聴覚重み付けのためのデータを求め、この重み付けのためのデータが端子307を介して、1段目の第2の符号化部120<sub>1</sub>の聴覚重み付け合成フィルタ312に送られる。また、当該聴覚重み付けフィルタ304では、図17のステップS2に示すように、入力音声データと量子化前の $\alpha$ パ

ラメータとから、前記聴覚重み付けした信号（前記図3の聴覚重み付けフィルタ125からの出力と同じ信号）を生成する。すなわち、まず、量子化前の $\alpha$ パラメータから聴覚重み付けフィルタ関数 $W(z)$ を生成し、さらに入力音声データ $x$ に当該フィルタ関数 $W(z)$ を適用して $x_s$ を生成し、これを上記聴覚重み付けした信号として、端子306を介して1段目の第2の符号化部120<sub>1</sub>の減算器313に送る。

【0357】1段目の第2の符号化部120<sub>1</sub>では、9ビットシェイプインデックス出力のストキャスティックコードブック（stochastic code book）310からの代表値出力（無声音のLPC残差に相当するノイズ出力）がゲイン回路311に送られ、このゲイン回路311にて、ストキャスティックコードブック310からの代表値出力に6ビットゲインインデックス出力のゲインコードブック315からのゲイン（スカラー値）を乗じ、このゲイン回路311にてゲインが乗じられた代表値出力が、 $1/A(z) = (1/H(z)) \cdot W(z)$ の聴覚重み付きの合成フィルタ312に送られる。この重み付きの合成フィルタ312からは、図17のステップS3のように、 $1/A(z)$ のゼロ入力応答出力が減算器313に送られる。当該減算器313では、上記聴覚重み付き合成フィルタ312からのゼロ入力応答出力と、上記聴覚重み付けフィルタ304からの上記聴覚重み付けした信号 $x_s$ とを用いた減算が行われ、この差分或いは誤差が参照ベクトル $r$ として取り出される。図17のステップS4に示すように、1段目の第2の符号化部120<sub>1</sub>でのサーチ時には、この参照ベクトル $r$ が、距離計算回路314に送られ、ここで距離計算が行われ、量子化誤差エネルギー $E$ を最小にするシェイプベクトル $s$ とゲイン $g$ がサーチされる。ただし、ここでの $1/A(z)$ はゼロ状態である。すなわち、コードブック中のシェイプベクトル $s$ をゼロ状態の $1/A(z)$ で合成したものを $s_{vn}$ とすると、式(40)を最小にするシェイプベクトル $s$ とゲイン $g$ をサーチする。

【0358】

【数79】

$$E = \sum_{n=0}^{N-1} (r(n) - gs_{yn}(n))^2 \quad \dots(40)$$

【0359】ここで、量子化誤差エネルギー $E$ を最小とする $s$ と $g$ をフルサーチしてもよいが、計算量を減らすために、以下のような方法をとることができる。なお、 $r(n)$ 等は、ベクトル $r$ 等の要素を表している。

【0360】第1の方法として、以下の式(41)に定義する $E_s$ を最小とするシェイプベクトル $s$ をサーチする。

【0361】

【数80】

$$E_s = \frac{\sum_{n=0}^{N-1} r(n)s_{yn}(n)}{\sqrt{\sum_{n=0}^{N-1} s_{yn}(n)^2}} \quad \dots(41)$$

【0362】第2の方法として、第1の方法により得られた $s$ より、理想的なゲインは、式(42)のようになるから、式(43)を最小とする $g$ をサーチする。

【0363】

【数81】

$$g_{ref} = \frac{\sum_{n=0}^{N-1} r(n)s_{yn}(n)}{\sum_{n=0}^{N-1} s_{yn}(n)^2} \quad \dots(42)$$

【0364】

$$E_g = (g_{ref} - g)^2 \quad (43)$$

ここで、 $E$ は $g$ の二次関数であるから、 $E_g$ を最小にする $g$ は $E$ を最小化する。

【0365】上記第1、第2の方法によって得られた $s$ と $g$ より、量子化誤差ベクトル $e$ は次の式(44)のように計算できる。

【0366】

$$e = r - gs_{vn} \quad (44)$$

これを、2段目の第2の符号化部120<sub>2</sub>のリファレンス入力として1段目と同様にして量子化する。

【0367】すなわち、上記1段目の第2の符号化部120<sub>1</sub>の聴覚重み付き合成フィルタ312からは、端子305及び端子307に供給された信号がそのまま2段目の第2の符号化部120<sub>2</sub>の聴覚重み付き合成フィルタ322に送られる。また、当該2段目の第2の符号化部120<sub>2</sub>減算器323には、1段目の第2の符号化部120<sub>1</sub>にて求めた上記量子化誤差ベクトル $e$ が供給される。

【0368】次に、図17のステップS5において、当該2段目の第2の符号化部120<sub>2</sub>でも1段目と同様に処理が行われる。すなわち、5ビットシェイプインデックス出力のストキャスティックコードブック320からの代表値出力がゲイン回路321に送られ、このゲイン回路321にて、当該コードブック320からの代表値出力に3ビットゲインインデックス出力のゲインコードブック325からのゲインを乗じ、このゲイン回路321の出力が、聴覚重み付きの合成フィルタ322に送られる。当該重み付きの合成フィルタ322からの出力は減算器323に送られ、当該減算器323にて上記聴覚重み付き合成フィルタ322からの出力と1段目の量子化誤差ベクトル $e$ との差分が求められ、この差分が距離計算回路324に送られてここで距離計算が行われ、量子化誤差エネルギー $E$ を最小にするシェイプベクトル $s$ とゲイン $g$ がサーチされる。



【0369】上述したような1段目の第2の符号化部120<sub>1</sub>のストキャストコードブック310からのシェイプインデックス出力及びゲインコードブック315からのゲインインデックス出力と、2段目の第2の符号化部120<sub>2</sub>のストキャストコードブック320からのインデックス出力及びゲインコードブック325からのインデックス出力は、インデックス出力切り換え回路330に送られるようになっている。ここで、当該第2の符号化部120から23ビット出力を行うときには、上記1段目と2段目の第2の符号化部120<sub>1</sub>及び120<sub>2</sub>のストキャストコードブック310、320及びゲインコードブック315、325からの各インデックスを合わせて出力し、一方、15ビット出力を行うときには、上記1段目の第2の符号化部120<sub>1</sub>のストキャストコードブック310とゲインコードブック315からの各インデックスを出力する。

【0370】その後は、ステップS6のようにフィルタ状態がアップデートされる。

【0371】ところで、本実施の形態では、2段目の第2の符号化部120<sub>2</sub>のインデックスビット数が、シェイプベクトルについては5ビットで、ゲインについては3ビットと非常に少ない。このような場合、適切なシェイプ、ゲインがコードブックに存在しないと、量子化誤差を減らすどころか逆に増やしてしまう可能性がある。

【0372】この問題を防ぐためには、ゲインに0を用意しておけばよいが、ゲインは3ビットしかなく、そのうちの一つを0にしてしまうのは量子化器の性能を大きく低下させてしまう。そこで、比較的多いビット数を割り当てたシェイプベクトルに、要素が全て0のベクトルを用意する。そして、このゼロベクトルを除いて、前述のサーチを行い、量子化誤差が最終的に増えてしまった場合に、ゼロベクトルを選択するようにする。なお、このときのゲインは任意である。これにより、2段目の第2の符号化部120<sub>2</sub>が量子化誤差を増すことを防ぐことができる。

【0373】なお、図16の例では、2段構成の場合を例に挙げているが、2段に限らず複数段構成とすることができる。この場合、1段目のクロズドループサーチによるベクトル量子化が終了したら、N段目(2≤N)ではN-1段目の量子化誤差をリファレンス入力として量子化を行い、さらにその量子化誤差をN+1段目のリファレンス入力とする。

【0374】上述したように、図16及び図17から、第2の符号化部に多段のベクトル量子化器を用いることにより、従来のような同じビット数のストレートベクトル量子化や共役コードブックなどを用いたものと比較して、計算量が少なくなる。特に、CELP符号化では、合成による分析(Analysis by Synthesis)法を用いたクロズドループサーチを用いた時間軸波形のベクトル量子化を行っているため、サーチの回数が少ないことが

重要である。また、2段の第2の符号化部120<sub>1</sub>と120<sub>2</sub>の両インデックス出力を用いる場合と、1段目の第2の符号化部120<sub>1</sub>のインデックス出力のみを用いる(2段目の第2の符号化部120<sub>2</sub>の出力インデックスを用いない)場合とを切り換えることにより、簡単にビット数を切り換えることが可能となっている。さらに上述したように、1段目と2段目の第2の符号化部120<sub>1</sub>と120<sub>2</sub>の両インデックス出力を合わせて出力するようなことを行えば、後のデコーダ側において例えば何れかを選ぶようにすることで、デコーダ側でも容易に対応できることになる。すなわち例えば6kbp/sでエンコードしたパラメータを、2kbp/sのデコーダでデコードするときに、デコーダ側で容易に対応できることになる。またさらに、例えば2段目の第2の符号化部120<sub>2</sub>のシェイプコードブックにゼロベクトルを含ませることにより、割り当てられたビット数が少ない場合でも、ゲインに0を加えるよりは少ない性能劣化で量子化誤差が増加することを防ぐことが可能となっている。

【0375】次に、上記ストキャストコードブックのコードベクトル(シェイプベクトル)は例えば以下のようにして生成することができる。

【0376】例えば、ストキャストコードブックのコードベクトルは、いわゆるガウシアンノイズのクリッピングにより生成することができる。具体的には、ガウシアンノイズを発生させ、これを適当なスレシホールド値でクリッピングし、それを正規化することで、コードブックを構成することができる。

【0377】ところが、音声には様々な形態があり、例えば「さ、し、す、せ、そ」のようなノイズに近い子音の音声には、ガウシアンノイズが適しているが、例えば「ば、び、ぶ、べ、ぼ」のような立ち上がりの激しい子音(急峻な子音)の音声については、対応しきれない。

【0378】そこで、本発明では、全コードベクトルのうち、適当な数はガウシアンノイズとし、残りを学習により求めて上記立ち上がりの激しい子音とノイズに近い子音の何れにも対応できるようにする。例えば、スレシホールド値を大きくとると、大きなピークを幾つか持つようなベクトルが得られ、一方、スレシホールド値を小さくすると、ガウシアンノイズそのものに近くなる。したがって、このようにクリッピングスレシホールド値のバリエーションを増やすことにより、例えば「ば、び、ぶ、べ、ぼ」のような立ち上がりの激しい子音や、例えば「さ、し、す、せ、そ」のようなノイズに近い子音などに対応でき、明瞭度を向上させることができるようになる。なお、図18には、図中実線で示すガウシアンノイズと図中点線で示すクリッピング後のノイズの様子を示している。また、図18の(A)はクリッピングスレシホールド値が1.0の場合(すなわちスレシホールド値が大きい場合)を、図18の(B)にはクリッピングスレシホールド値が0.4の場合(すなわちスレシホー

ルド値が小さい場合)を示している。この図18の(A)及び(B)から、スレシホールド値を大きくすると、大きなピークを幾つか持つようなベクトルが得られ、一方、スレシホールド値を小さくすると、ガウシアンノイズそのものに近くなることが判る。

【0379】このようなことを実現するため、先ず、ガウシアンノイズのクリッピングにより初期コードブックを構成し、さらに予め適当な数だけ学習を行わないコードベクトルを決めておく。この学習しないコードベクトルは、その分散値が小さいものから順に選ぶようにする。これは、例えば「さ、し、す、せ、そ」のようなノイズに近い子音に対応させるためである。一方、学習を行って求めるコードベクトルは、当該学習のアルゴリズムとしてLBGアルゴリズムを用いるようにする。ここで最適エンコード条件(Nearest Neighbour Condition)でのエンコードは固定したコードベクトルと、学習対象のコードベクトル両方を使用して行う。セントロイドコンディション(Centroid Condition)においては、学習対象のコードベクトルのみをアップデートする。これにより、学習対象となったコードベクトルは「ば、び、ぶ、べ、ぼ」などの立ち上がりの激しい子音に対応するようになる。

【0380】なお、ゲインは通常通りの学習を行うことで、これらのコードベクトルに対して最適なものが学習できる。

【0381】上述したガウシアンノイズのクリッピングによるコードブックの構成のための処理の流れを図19に示す。

【0382】この図19において、ステップS10では、初期化として、学習回数 $n=0$ とし、誤差 $D_0=\infty$ とし、最大学習回数 $n_{max}$ を決定し、学習終了条件を決めるスレシホールド値 $\varepsilon$ を決定する。

【0383】次のステップS11では、ガウシアンノイズのクリッピングによる初期コードブックを生成し、ステップS12では学習を行わないコードベクトルとして一部のコードベクトルを固定する。

【0384】次にステップS13では上記コードブックを用いてエンコードを行い、ステップS14では誤差を算出し、ステップS15では $(D_{n-1}-D_n)/D_n < \varepsilon$ 、若しくは $n=n_{max}$ かを判断し、Yesと判断した場合には処理を終了し、Noと判断した場合にはステップS16に進む。

【0385】ステップS16ではエンコードに使用されなかったコードベクトルの処理を行い、次のステップS17ではコードブックのアップデートを行う。次にステップS18では学習回数 $n$ を1インクリメントし、その後ステップS13に戻る。

【0386】次に、図3の音声信号符号化装置において、V/UV(有声音/無声音)判定部115の具体例について説明する。

【0387】このV/UV判定部115においては、直交変換回路145からの出力と、高精度ピッチサーチ部146からの最適ピッチと、スペクトル評価部148からのスペクトル振幅データと、オープンループピッチサーチ部141からの正規化自己相関最大値 $r(p)$ と、ゼロクロスカウンタ412からのゼロクロスカウント値とに基づいて、当該フレームのV/UV判定が行われる。さらに、MBEの場合と同様な各バンド毎のV/UV判定結果の境界位置も当該フレームのV/UV判定の条件としている。

【0388】このMBEの場合の各バンド毎のV/UV判定結果を用いたV/UV判定条件について以下に説明する。

【0389】MBEの場合の第 $m$ 番目のハーモニクスの大きさを表すパラメータあるいは振幅 $|A_m|$ は、

【0390】

【数82】

$$\therefore |A_m| = \frac{\sum_{j=a_m}^{b_m} |S(j)| |E(j)|}{\sum_{j=a_m}^{b_m} |E(j)|^2}$$

【0391】により表せる。この式において、 $|S(j)|$ は、LPC残差をDFTしたスペクトルであり、 $|E(j)|$ は、基底信号のスペクトル、具体的には256ポイントのハミング窓をDFTしたものである。また、 $a_m$ 及び $b_m$ は、第 $m$ 番目のハーモニクスに対応する第 $m$ バンドに対応する周波数をインデックス $j$ で表現したときの下限値及び上限値である。また、各バンド毎のV/UV判定のために、NSR(ノイズtoシグナル比)を利用する。この第 $m$ バンドのNSRは、

【0392】

【数83】

$$NSR = \frac{\sum_{j=a_m}^{b_m} \{ |S(j)| - |A_m| |E(j)| \}^2}{\sum_{j=a_m}^{b_m} |S(j)|^2}$$

【0393】と表せ、このNSR値が所定の閾値(例えば0.3)より大のとき(エラーが大きい)ときには、そのバンドでの $|A_m| |E(j)|$ による $|S(j)|$ の近似が良くない(上記励起信号 $|E(j)|$ が基底として不適当である)と判断でき、当該バンドをUV(Unvoiced、無声音)と判別する。これ以外のときは、近似がある程度良好に行われていると判断でき、そのバンドをV(Voiced、有声音)と判別する。

【0394】ここで、上記各バンド(ハーモニクス)のNSRは、各ハーモニクス毎のスペクトル類似度をあらわしている。NSRのハーモニクスのゲインによる重み付け和をとったものを $NSR_{all}$ として次のように定義する。

$$【0395】NSR_{s11} = (\Sigma_n |A_n| NSR_n) / (\Sigma_n |A_n|)$$

このスペクトル類似度 $NSR_{s11}$ がある閾値より大きい  
か小さいかにより、V/UV判定に用いるルールベース  
を決定する。ここでは、この閾値を $Th_{NSR} = 0.3$ とし  
ておく。このルールベースは、フレームパワー、ゼロク  
ロス、LPC残差の自己相関の最大値に関するものであ  
り、 $NSR_{s11} < Th_{NSR}$ のときに用いられるルールベ  
ースでは、ルールが適用されるとVとなり適用されるル  
ールがなかった場合はUVとなる。

【0396】また、 $NSR_{s11} \geq Th_{NSR}$ のときに用い  
られるルールベースでは、ルールが適用されるとUV、  
適用されないとVとなる。

【0397】ここで、具体的なルールは、次のようなも  
のである。 $NSR_{s11} < Th_{NSR}$ のとき、

if numZeroXP < 24, & frmPow > 340, & r0 > 0.32  
then V

$NSR_{s11} \geq Th_{NSR}$ のとき、

if numZeroXP > 30, & frmPow < 900, & r0 < 0.23  
then UV

ただし、各変数は次のように定義される。

numZeroXP: 1フレーム当たりのゼロクロス回数

frmPow: フレームパワー

r0: 自己相関最大値

上記のようなルールの集合であるルールに照合すること  
で、V/UVを判定する。

【0398】次に、図4の音声復号化装置（デコーダ）  
の要部のより具体的な構成及び動作について説明する。

【0399】スペクトルエンベロープの逆ベクトル量子  
化器212においては、上述したような音声符号化装置  
（エンコーダ）側でのベクトル量子化器の構成に対応し  
た逆ベクトル量子化構成が用いられる。

【0400】例えば、エンコーダ側で上記図10に示し  
た構成によりベクトル量子化が施されている場合に、デ  
コーダ側では、与えられたインデクスに応じて符号帳5  
30のシェイプコードブックCB0、CB1及びゲイン\*

\*コードブックDB<sub>0</sub>から、コードベクトル $s_0$ 、 $s_1$ 及び  
ゲイン $g$ がそれぞれ読み出され、 $g(s_0 + s_1)$ の固定  
次元（例えば44次元）のベクトルとして取り出され、  
元のハーモニクススペクトルのベクトルの次元数に対応  
する可変次元のベクトルに変換（固定/可変次元変換）  
される。

【0401】また、エンコーダ側で、図12～図15の  
ように、可変次元ベクトルに固定次元コードベクトルを  
加算するようなベクトル量子化器の構成を有する場合に  
は、デコーダ側では、可変次元用のコードブック（例  
えば図12のコードブックCB0）から読み出されたコード  
ベクトルについては固定/可変次元変換し、これに固  
定次元用のコードブック（図12ではコードブックCB  
1）から読み出された固定次元のコードベクトルをハー  
モニクスの低域側から次元数分だけ加算して、取り出す  
ようにしている。

【0402】次に、図4のLPC合成フィルタ214  
は、上述したように、V（有声音）用の合成フィルタ2  
36と、UV（無声音）用の合成フィルタ237とに分  
離されている。すなわち、合成フィルタを分離せずにV  
/UVの区別なしに連続的にLSPの補間を20サンプ  
ルすなわち2.5 msec 毎に行う場合には、V→UV、  
UV→Vの遷移（トランジェント）部において、全く性  
質の異なるLSP同士を補間することになり、Vの残差  
にUVのLPCが、UVの残差にVのLPCが用いられ  
ることにより異音が発生するが、このような悪影響を防  
止するために、LPC合成フィルタをV用とUV用とで  
分離し、LPCの係数補間をVとUVとで独立に行わせ  
たものである。

【0403】この場合の、LPC合成フィルタ236、  
237の係数補間方法について説明する。これは、次の  
表3に示すように、V/UVの状態に応じてLSPの補  
間を切り換えている。

【0404】

【表3】

	Hv(z)		Huv(z)	
	前フレーム	現フレーム	前フレーム	現フレーム
V→V	伝送された LSP	伝送された LSP	均等間隔 LSP	均等間隔 LSP
V→UV	伝送された LSP	均等間隔 LSP	均等間隔 LSP	伝送された LSP
UV→V	均等間隔 LSP	伝送された LSP	伝送された LSP	均等間隔 LSP
UV→UV	均等間隔 LSP	均等間隔 LSP	伝送された LSP	伝送された LSP

【0405】この表3において、均等間隔LSPとは、  
例えば10次のLPC分析の例で述べると、フィルタの  
特性がフラットでゲインが1のときの $\alpha$ パラメータ、す  
なわち  $\alpha_0 = 1$ ,  $\alpha_1 = \alpha_2 = \dots = \alpha_{10} = 0$ に対応す

るLSPであり、

$$LSP_i = (\pi/11) \times i \quad 0 \leq i \leq 10$$

である。

【0406】このような10次のLPC分析、すなわち

10次のLSPの場合は、図20に示す通り、 $0 \sim \pi$ の間を11等分した位置に均等間隔で配置されたLSPで、完全にフラットなスペクトルに対応している。合成フィルタの全帯域ゲインはこのときが最小のスルー特性となる。

【0407】図21は、ゲイン変化の様子を概略的に示す図であり、UV（無声音）部分からV（有声音）部分への遷移時における $1/H_{uv}(z)$ のゲイン及び $1/H_v(z)$ のゲインの変化の様子を示している。

【0408】ここで、補間を行う単位は、フレーム間隔が160サンプル（20msec）のとき、 $1/H_v(z)$ の係数は2.5msec（20サンプル）毎、また $1/H_{uv}(z)$ の係数は、ビットレートが2kbpsで10msec（80サンプル）、6kbpsで5msec（40サンプル）毎である。なお、UV時はエンコード側の第2の符号化部120で合成による分析法を用いた波形マッチングを行っているので、必ずしも均等間隔LSPと補間せずとも、隣接するV部分のLSPとの補間を行ってもよい。ここで、第2の符号化部120におけるUV部の符号化処理においては、 $V \rightarrow UV$ への遷移部で $1/A(z)$ の重み付き合成フィルタ122の内部状態をクリアすることによりゼロインプットレスポンスを0にする。

【0409】これらのLPC合成フィルタ236、237からの出力は、それぞれ独立に設けられたポストフィルタ238v、238uに送られており、ポストフィルタもVとUVとで独立にかけることにより、ポストフィルタの強度、周波数特性をVとUVとで異なる値に設定している。

【0410】次に、LPC残差信号、すなわちLPC合成フィルタ入力であるエクサイテーションの、V部とUV部のつなぎ部分の窓かけについて説明する。これは、図4の有声音合成部211のサイン波合成回路215と、無声音合成部220の窓かけ回路223とによりそれぞれ行われるものである。なお、エクサイテーションのV部の合成方法については、本件出願人が先に提案した特願平4-91422号の明細書及び図面に具体的な説明が、また、V部の高速合成方法については、本件出願人が先に提案した特願平6-198451号の明細書及び図面に具体的な説明が、それぞれ開示されている。今回の具体例では、この高速合成方法を用いてV部のエクサイテーションを生成している。

【0411】V（有声音）部分では、隣接するフレームのスペクトルを用いてスペクトルを補間してサイン波合成するため、図22に示すように、第nフレームと第n+1フレームとの間にかかる全ての波形を作ることができる。しかし、図22の第n+1フレームと第n+2フレームとのように、VとUV（無声音）に跨る部分、あるいはその逆の部分では、UV部分は、フレーム中に±80サンプル（全160サンプル=1フレーム間隔）のデータのみをエンコード及びデコードしている。このた

め、図23に示すように、V側ではフレームとフレームとの間の中心点CNを越えて窓かけを行い、UV側では中心点CN移行の窓かけを行って、接続部分をオーバーラップさせている。UV→Vの遷移（トランジェント）部分では、その逆を行っている。なお、V側の窓かけは破線のようにしてもよい。

【0412】次に、V（有声音）部分でのノイズ合成及びノイズ加算について説明する。これは、図4のノイズ合成回路216、重み付き重畳回路217、及び加算器218を用いて、有声音部分のLPC合成フィルタ入力となるエクサイテーションについて、次のパラメータを考慮したノイズをLPC残差信号の有声音部分に加えることにより行われる。

【0413】すなわち、上記パラメータとしては、ピッチラグPch、有声音のスペクトル振幅 $Am[i]$ 、フレーム内の最大スペクトル振幅 $Am_{max}$ 、及び残差信号のレベル $Lev$ を挙げることができる。ここで、ピッチラグPchは、所定のサンプリング周波数 $f_s$ （例えば $f_s = 8kHz$ ）でのピッチ周期内のサンプル数であり、スペクトル振幅 $Am[i]$ のiは、 $f_s/2$ の帯域内でのハーモニックスの本数を $I = Pch/2$ とすると、 $0 < i < I$ の範囲内の整数である。

【0414】このノイズ合成回路216による処理は、例えばMBE（マルチバンド励起）符号化の無声音の合成と同様な方法で行われる。図24は、ノイズ合成回路216の具体例を示している。

【0415】すなわち図24において、ホワイトノイズ発生部401からは、時間軸上のホワイトノイズ信号波形に所定の長さ（例えば256サンプル）で適当な窓関数（例えばハミング窓）により窓かけされたガウシャノイズが出力され、これがSTFT処理部402によりSTFT（ショートタームフーリエ変換）処理を施すことにより、ノイズの周波数軸上のパワースペクトルを得る。このSTFT処理部402からのパワースペクトルを振幅処理のための乗算器403に送り、ノイズ振幅制御回路410からの出力を乗算している。乗算器403からの出力は、ISTFT処理部404に送られ、位相は元のホワイトノイズの位相を用いて逆STFT処理を施すことにより時間軸上の信号に変換する。ISTFT処理部404からの出力は、重み付き重畳加算回路217に送られる。

【0416】なお、上記図24の例においては、ホワイトノイズ発生部401から時間領域のノイズを発生してそれをSTFT等の直交変換を行うことで周波数領域のノイズを得ていたが、ノイズ発生部から直接的に周波数領域のノイズを発生するようにしてもよい。すなわち、周波数領域のパラメータを直接発生することにより、STFTやFFT等の直交変換処理が節約できる。

【0417】具体的には、 $\pm x$ の範囲の乱数を発生しそれをFFTスペクトルの実部と虚部として扱うようにす

る方法や、0から最大値(max)までの範囲の正の乱数を発生しそれをFFTスペクトルの振幅として扱い、 $-\pi$ から $\pi$ までの乱数を発生しそれをFFTスペクトルの位相として扱う方法などが挙げられる。

【0418】こうすることにより、図24のSTFT処理部402が不要となり、構成の簡略化あるいは演算量の低減が図れる。

【0419】ノイズ振幅制御回路410は、例えば図25のような基本構成を有し、上記図4のスペクトルエンベロープの逆量子化器212から端子411を介して与えられるV(有声音)についての上記スペクトル振幅Am[i]と、上記図4の入力端子204から端子412を介して与えられる上記ピッチラグPchに基づいて、乗算器403での乗算係数を制御することにより、合成される\*

$$\begin{aligned} f_1(Pch, Am[i]) &= 0 & (0 < i < Noise\_b \times I) \\ f_1(Pch, Am[i]) &= Am[i] \times noise\_mix & (Noise\_b \times I \leq i < I) \\ noise\_mix &= K \times Pch / 2.0 \end{aligned}$$

とすることが挙げられる。

【0422】ただし、noise\_mixの最大値は、noise\_mix\_maxとし、その値でクリップする。一例として、K=0.02、noise\_mix\_max=0.3、Noise\_b=0.7とすることが挙げられる。ここで、Noise\_bは、全帯域の割合からこのノイズの付加を行うかを定める定数である。本例では、7割より高域側、すなわちfs=8kHzのとき、4000×0.7=2800Hzから4000Hzの間でノイズ

$$\begin{aligned} f_2(Pch, Am[i], Amax) &= 0 & (0 < i < Noise\_b \times I) \\ f_2(Pch, Am[i], Amax) &= Am[i] \times noise\_mix & (Noise\_b \times I \leq i < I) \\ noise\_mix &= K \times Pch / 2.0 \end{aligned}$$

を挙げることができる。ただし、noise\_mixの最大値は、noise\_mix\_maxとし、一例として、K=0.02、noise\_mix\_max=0.3、Noise\_b=0.7とすることが挙げられる。

【0425】さらに、もしAm[i]×noise\_mix>Amax×C×noise\_mixならば、

$f_2(Pch, Am[i], Amax) = Amax \times C \times noise\_mix$ とする。ここで、定数Cは、C=0.3としている。この条件式によりノイズレベルが大きくなり過ぎることを防止できるため、上記K、noise\_mix\_maxをさらに大きくしてもよく、高域のレベルも比較的大きいときにノイズレベルを高めることができる。

【0426】次に、ノイズ合成加算の第3の具体例として、上記ノイズ振幅Am\_noise[i]を、上記4つのパラメータの内の4つ全ての関数 $f_3(Pch, Am[i], Amax, Lev)$ とすることもできる。

【0427】このような関数 $f_3(Pch, Am[i], Amax, Lev)$ の具体例は、基本的には上記第2の具体例の関数 $f_2(Pch, Am[i], Amax)$ と同様である。ただし、残差信号レベルLevは、スペクトル振幅Am[i]のrms(root mean square)、あるいは時間軸上で測定した信号レベルである。上記第2の具体例との違いは、Kの値とnoise\_

\*ノイズ振幅Am\_noise[i]を求めている。すなわち図25において、スペクトル振幅Am[i]とピッチラグPchが入力される最適なnoise\_mix値の算出回路416からの出力をノイズの重み付け回路417で重み付けし、得られた出力を乗算器418に送ってスペクトル振幅Am[i]と乗算することにより、ノイズ振幅Am\_noise[i]を得ている。

【0420】ここで、ノイズ合成加算の第1の具体例として、ノイズ振幅Am\_noise[i]が、上記4つのパラメータの内の2つ、すなわちピッチラグPch及びスペクトル振幅Am[i]の関数 $f_1(Pch, Am[i])$ となる場合について説明する。

【0421】このような関数 $f_1(Pch, Am[i])$ の具体例として、

※イズを付加するようにしている。

【0423】次に、ノイズ合成加算の第2の具体例として、上記ノイズ振幅Am\_noise[i]を、上記4つのパラメータの内の3つ、すなわちピッチラグPch、スペクトル振幅Am[i]及び最大スペクトル振幅Amaxの関数 $f_2(Pch, Am[i], Amax)$ とする場合について説明する。

【0424】このような関数 $f_2(Pch, Am[i], Amax)$ の具体例として、

mix\_maxの値とをLevの関数とする点である。すなわち、Levが小さくなったときには、K、noise\_mix\_maxの各値を大きめに設定し、Levが大きいたときは小さめに設定する。あるいは、連続的にLevの値を逆比例させてもよい。

【0428】次に、ポストフィルタ238v、238uについて説明する。

【0429】図26は、図4の例のポストフィルタ238v、238uとして用いられるポストフィルタを示しており、ポストフィルタの要部となるスペクトル整形フィルタ440は、ホルマント強調フィルタ441と高域強調フィルタ442とから成っている。このスペクトル整形フィルタ440からの出力は、スペクトル整形によるゲイン変化を補正するためのゲイン調整回路443に送られており、このゲイン調整回路443のゲインGは、ゲイン制御回路445により、スペクトル整形フィルタ440の入力xと出力yと比較してゲイン変化を計算し、補正値を算出することで決定される。

【0430】スペクトル整形フィルタの440特性PF(z)は、LPC合成フィルタの分母Hv(z)、Hu(z)の係数、いわゆる $\alpha$ パラメータを $\alpha_i$ とすると、

【0431】

【数84】

$$PF(z) = \frac{\sum_{i=0}^P \alpha_i \beta^i z^{-i}}{\sum_{i=0}^P \alpha_i \gamma^i z^{-i}} (1 - k z^{-1})$$

【0432】と表せる。この式の分数部分がホルマント強調フィルタ特性を、 $(1 - k z^{-1})$ の部分が高域強調フィルタ特性をそれぞれ表す。また、 $\beta$ 、 $\gamma$ 、 $k$ は定数であり、一例として $\beta = 0.6$ 、 $\gamma = 0.8$ 、 $k = 0.3$ を挙げることができる。

【0433】また、ゲイン調整回路443のゲイン $G$ は、

【0434】

【数85】

$$G = \sqrt{\frac{\sum_{i=0}^{159} x^2(i)}{\sum_{i=0}^{159} y^2(i)}}$$

【0435】としている。この式中の $x(i)$ はスペクトル整形フィルタ440の入力、 $y(i)$ はスペクトル整形フィルタ440の出力である。

【0436】ここで、上記スペクトル整形フィルタ440の係数の更新周期は、図27に示すように、LPC合成フィルタの係数である $\alpha$ パラメータの更新周期と同じく20サンプル、2.5msecであるのに対して、ゲイン調整回路443のゲイン $G$ の更新周期は、160サンプル、20msecである。

【0437】このように、ポストフィルタのスペクトル整形フィルタ440の係数の更新周期に比較して、ゲイン調整回路443のゲイン $G$ の更新周期を長くすることにより、ゲイン調整の変動による悪影響を防止している。

【0438】すなわち、一般のポストフィルタにおいては、スペクトル整形フィルタの係数の更新周期とゲインの更新周期とを同じにしており、このとき、ゲインの更新周期を20サンプル、2.5msecとすると、図27からも明らかなように、1ピッチ周期の中で変動することになり、クリックノイズを生じる原因となる。そこで本例においては、ゲインの切換周期をより長く、例えば1フレーム分の160サンプル、20msecとすることにより、急激なゲインの変動を防止することができる。また逆に、スペクトル整形フィルタの係数の更新周期を160サンプル、20msecとするときには、円滑なフィルタ特性の変化が得られず、合成波形に悪影響が生じるが、このフィルタ係数の更新周期を20サンプル、2.5msecと短くすることにより、効果的なポストフィルタ処理が可能となる。

【0439】なお、隣接するフレーム間でのゲインのつなぎ処理は、図28に示すように、前フレームのフィルタ係数及びゲインと、現フレームのフィルタ係数及びゲインとを用いて算出した結果に、次のような三角窓

$$W(i) = i/20 \quad (0 \leq i \leq 20)$$

と

$$1 - W(i) \quad (0 \leq i \leq 20)$$

をかけてフェードイン、フェードアウトを行って加算する。図28では、前フレームのゲイン $G_1$ が現フレームのゲイン $G_2$ に変化する様子を示している。すなわち、オーバーラップ部分では、前フレームのゲイン、フィルタ係数を使用する割合が徐々に減衰し、現フレームのゲイン、フィルタ係数の使用が徐々に増大する。なお、図28の時刻Tにおけるフィルタの内部状態は、現フレームのフィルタ、前フレームのフィルタ共に同じもの、すなわち前フレームの最終状態からスタートする。

【0440】以上説明したような信号符号化装置及び信号復号化装置は、例えば図29及び図30に示すような携帯通信端末あるいは携帯電話機等に使用される音声コーデックとして用いることができる。

【0441】すなわち、図29は、上記図1、図3に示したような構成を有する音声符号化部160を用いて成る携帯端末の送信側構成を示している。この図29のマイクロホン161で集音された音声信号は、アンプ162で増幅され、A/D（アナログ/デジタル）変換器163でデジタル信号に変換されて、音声符号化部160に送られる。この音声符号化部160は、上述した図1、図3に示すような構成を有しており、この入力端子101に上記A/D変換器163からのデジタル信号が入力される。音声符号化部160では、上記図1、図3と共に説明したような符号化処理が行われ、図1、図2の各出力端子からの出力信号は、音声符号化部160の出力信号として、伝送路符号化部164に送られる。伝送路符号化部164では、いわゆるチャネルコーディング処理が施され、その出力信号が変調回路165に送られて変調され、D/A（デジタル/アナログ）変換器166、RFアンプ167を介して、アンテナ168に送られる。

【0442】また、図30は、上記図2、図4に示したような構成を有する音声復号化部260を用いて成る携帯端末の受信側構成を示している。この図30のアンテナ261で受信された音声信号は、RFアンプ262で増幅され、A/D（アナログ/デジタル）変換器263を介して、復調回路264に送られ、復調信号が伝送路復号化部265に送られる。264からの出力信号は、上記図2、図4に示すような構成を有する音声復号化部260に送られる。音声復号化部260では、上記図2、図4と共に説明したような復号化処理が施され、図2、図4の出力端子201からの出力信号が、音声復号化部260からの信号としてD/A（デジタル/ア

ナログ)変換器266に送られる。このD/A変換器266からのアナログ音声信号がスピーカ268に送られる。

【0443】なお、本発明は上記実施の形態のみに限定されるものではなく、例えば上記図1、図3の音声分析側(エンコード側)の構成や、図2、図4の音声合成側(デコード側)の構成については、各部をハードウェア的に記載しているが、いわゆるDSP(デジタル信号プロセッサ)等を用いてソフトウェアプログラムにより実現することも可能である。また、ベクトル量子化は、音声符号化のみならず、他の種々の信号のベクトル量子化に適用できる。さらに、本発明の音声符号化方法や装置の適用範囲は、伝送や記録再生に限定されず、ピッチ変換やスピード変換、規則音声合成、あるいは雑音抑圧のような種々の用途に応用できることは勿論である。

【0444】

【発明の効果】以上の説明から明らかなように、本発明によれば、ベクトル量子化しようとする入力ベクトルと符号帳(コードブック)に蓄えられた全てのコードベクトルとの類似度を近似計算により求めて類似度の高い複数のコードベクトルを予備選択し、この予備選択により選択された複数のコードベクトルについて上記入力ベクトルとの誤差を最小化する最適のコードベクトルを本選択することにより、簡略化された演算の予備選択である程度候補を絞り込んだ後に、これらの候補についてのみ高精度の本選択を行わせることができ、符号帳検索(コードブックサーチ)に要する演算量を低減できる。

【0445】ここで上記符号帳は、複数の符号帳を組み合わせて構成され、最適の組み合わせを構成するコードベクトルが各符号帳毎にそれぞれ選択されることが挙げられる。また、上記類似度としては、上記入力ベクトルと上記コードベクトルとの内積、重み付き内積、又はこれらを各コードベクトルのノルム又は重み付きノルムで割ったもののいずれかを用いることができる。

【0446】このようなベクトル量子化を音声符号化に適用することができ、例えば、短期予測残差をサイン波分析してハーモニクススペクトルを求め、符号化単位毎の上記ハーモニクススペクトルに基づくパラメータを入力ベクトルとしてベクトル量子化する際に適用することができる。これによって、代表的な(平均の)重みによる重み付きノルムを各コードベクトル毎にオフラインで計算しておき、予備選択時に用いることで、より精度の高い予備選択が可能となり、予備選択を導入することによる音質劣化が防止できる。

【図面の簡単な説明】

【図1】本発明に係るベクトル量子化方法が適用された音声符号化方法の実施の形態となる音声符号化装置の基本構成を示すブロック図である。

【図2】図1の音声符号化装置により符号化された信号を復号化するための音声復号化装置の基本構成を示すブ

ロック図である。

【図3】本発明の実施の形態となる音声符号化装置のより具体的な構成を示すブロック図である。

【図4】図2の音声復号化装置のより具体的な構成を示すブロック図である。

【図5】LSP量子化部の基本構成を示すブロック図である。

【図6】LSP量子化部のより具体的な構成を示すブロック図である。

【図7】ベクトル量子化部の基本構成を示すブロック図である。

【図8】ベクトル量子化部のより具体的な構成を示すブロック図である。

【図9】重み付けの重みの具体例を示すグラフである。

【図10】符号帳検索を可変次元で行うベクトル量子化器の構成例を示すブロック回路図である。

【図11】符号帳検索を可変次元で行うベクトル量子化器の他の構成例を示すブロック回路図である。

【図12】可変次元用の符号帳と固定次元用の符号帳とを用いるベクトル量子化器の第1の構成例を示すブロック回路図である。

【図13】可変次元用の符号帳と固定次元用の符号帳とを用いるベクトル量子化器の第2の構成例を示すブロック回路図である。

【図14】可変次元用の符号帳と固定次元用の符号帳とを用いるベクトル量子化器の第3の構成例を示すブロック回路図である。

【図15】可変次元用の符号帳と固定次元用の符号帳とを用いるベクトル量子化器の第5の構成例を示すブロック回路図である。

【図16】本発明の音声信号符号化装置のCELP符号化部分(第2の符号化部)の具体的な構成を示すブロック回路図である。

【図17】図16の構成における処理の流れを示すフローチャートである。

【図18】ガウシアンノイズと、異なるスレシホールド値でのクリッピング後のノイズの様子を示す図である。

【図19】学習によってシェイプコードブックを生成する際の処理の流れを示すフローチャートである。

【図20】10次のLPC分析により得られた $\alpha$ パラメータに基づく10次のLSP(線スペクトル対)を示す図である。

【図21】UV(無声音)フレームからV(有声音)フレームへのゲイン変化の様子を説明するための図である。

【図22】フレーム毎に合成されるスペクトルや波形の補間処理を説明するための図である。

【図23】V(有声音)フレームとUV(無声音)フレームとの接続部でのオーバーラップを説明するための図である。



【図24】有声音合成の際のノイズ加算処理を説明するための図である。

【図25】有声音合成の際に加算されるノイズの振幅計算の例を示す図である。

【図26】ポストフィルタの構成例を示す図である。

【図27】ポストフィルタのフィルタ係数更新周期とゲイン更新周期とを説明するための図である。

【図28】ポストフィルタのゲイン、フィルタ係数のフレーム境界部分でのつなぎ処理を説明するための図である。

【図29】本発明の実施の形態となる音声信号符号化装置が用いられる携帯端末の送信側構成を示すブロック図である。

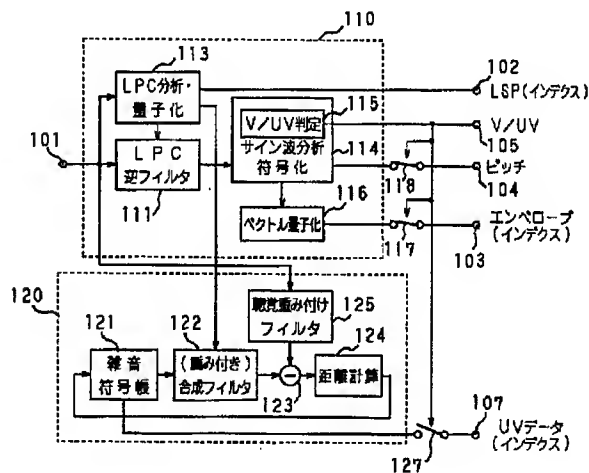
【図30】本発明の実施の形態となる音声信号復号化装置が用いられる携帯端末の受信側構成を示すブロック図\*

\*である。

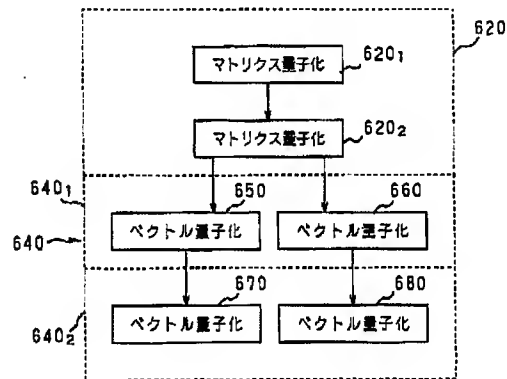
【符号の説明】

110 第1の符号化部、111 LPC逆フィルタ、113 LPC分析・量子化部、114 サイン波分析符号化部、115 V/UV判定部、116 ベクトル量子化器、120 第2の符号化部、121 雑音符号帳、122 重み付き合成フィルタ、123 減算器、124 距離計算回路、125 聴覚重み付けフィルタ、530 符号帳(コードブック)、531 シェイプコードブック、532 ゲインコードブック、533ゲイン回路、535 仮選択用の選択回路、542 可変/固定次元変換回路、544 固定/可変次元変換回路、545 本選択用の選択回路

【図1】

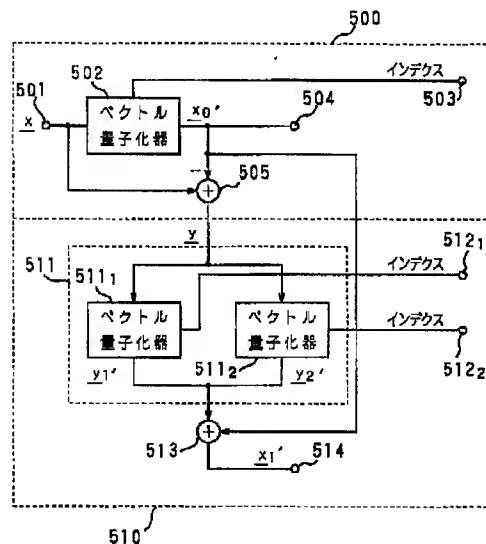
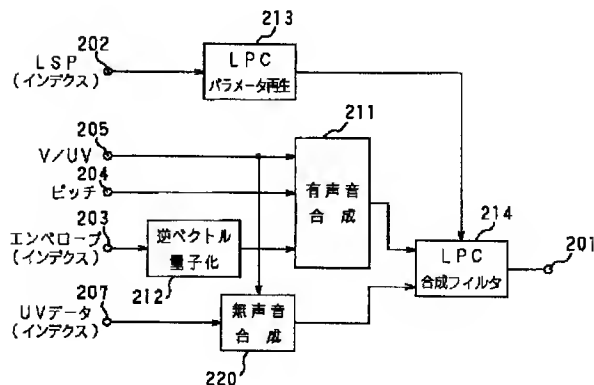


【図5】



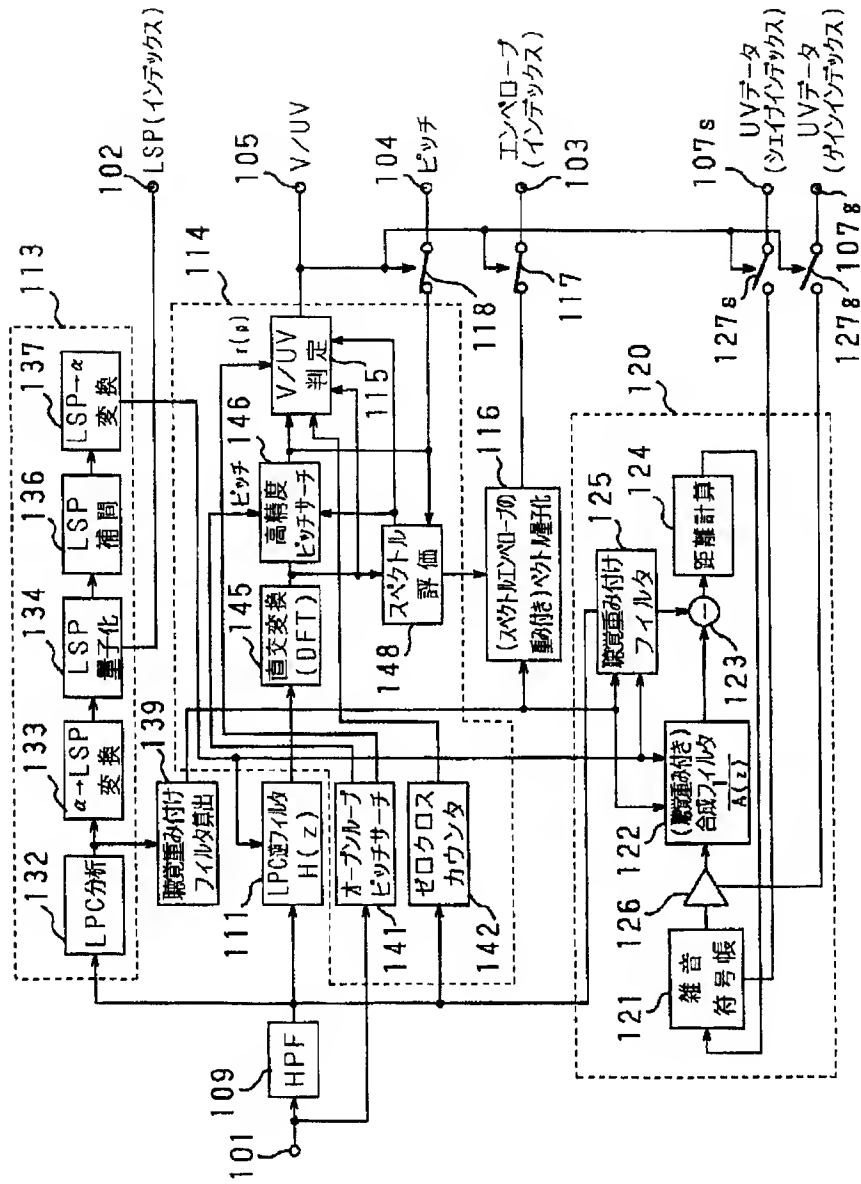
【図7】

【図2】

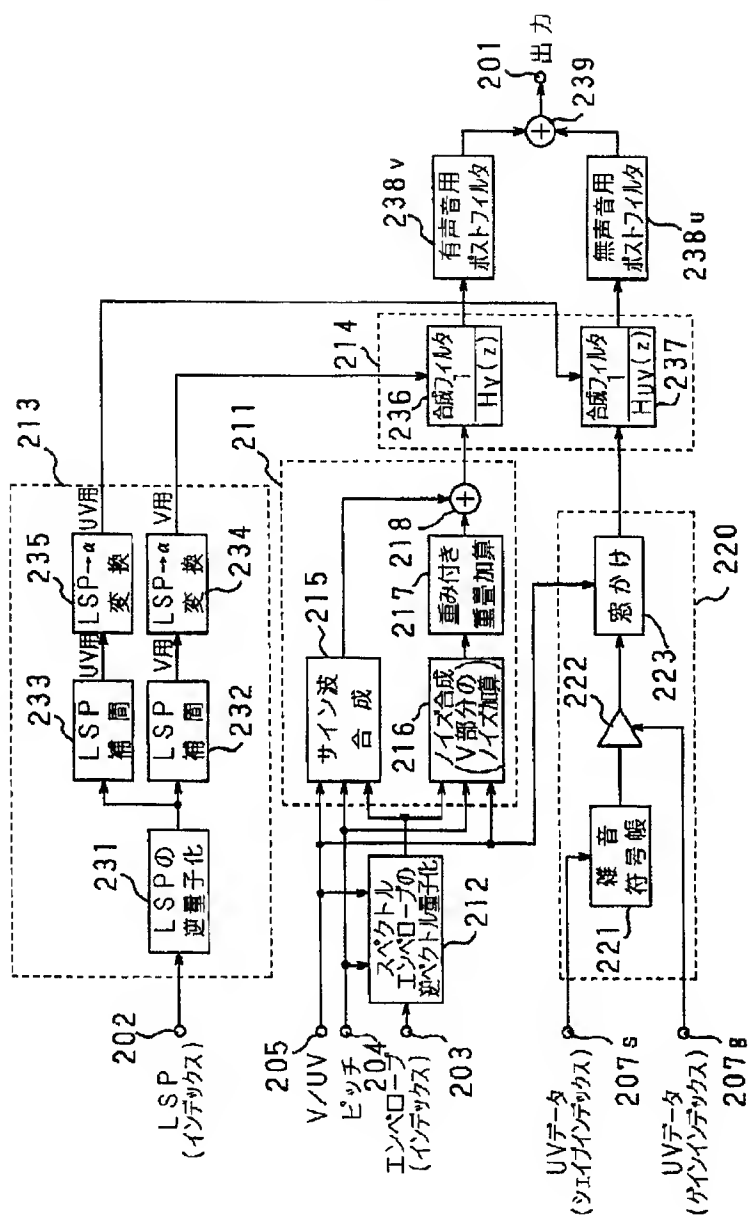




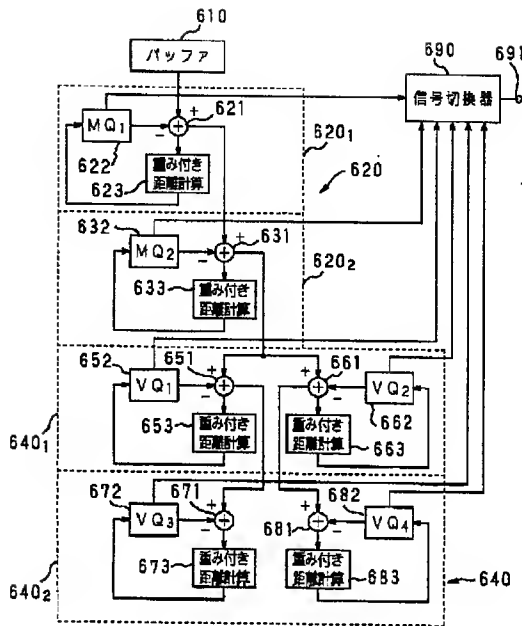
【図3】



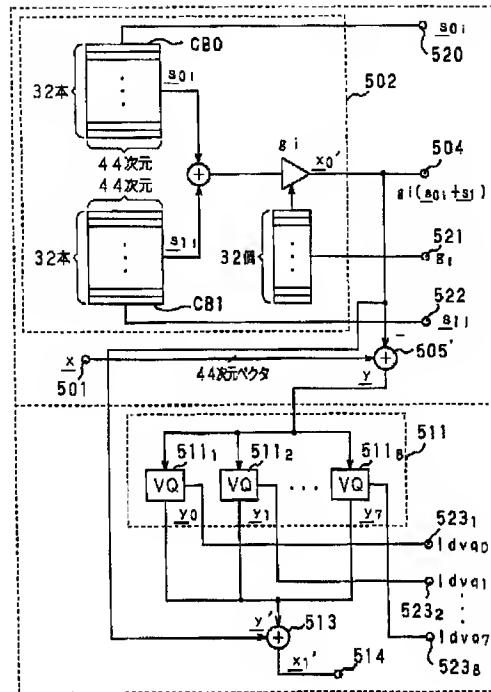
【図4】



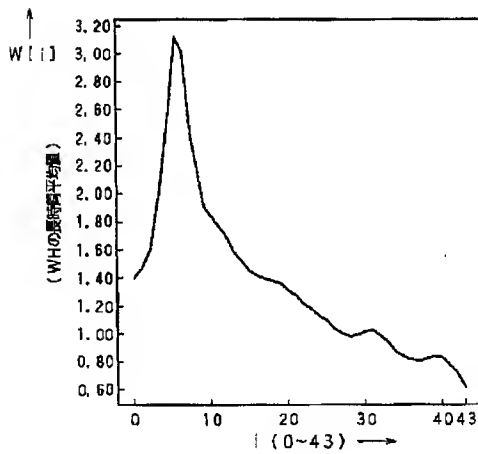
【図6】



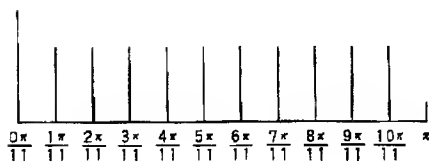
【図8】



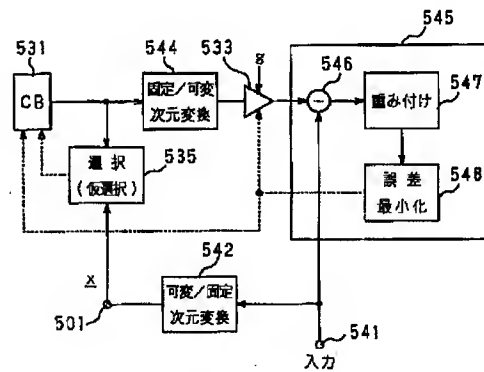
【図9】



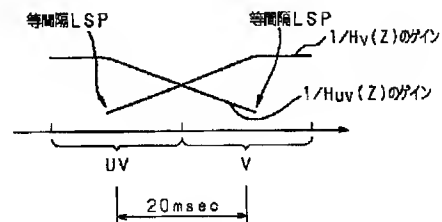
【図20】



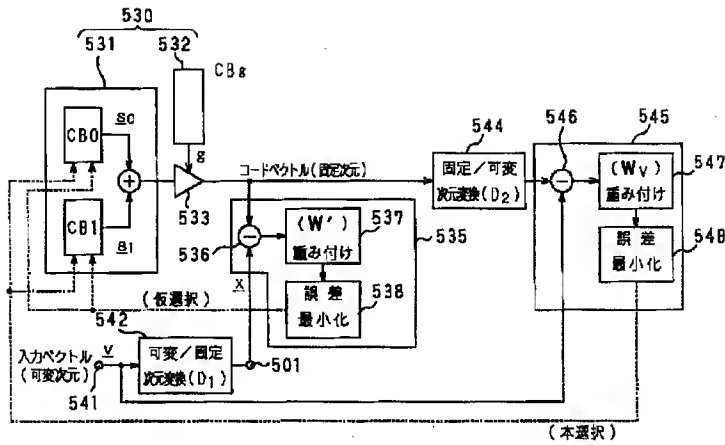
【図11】



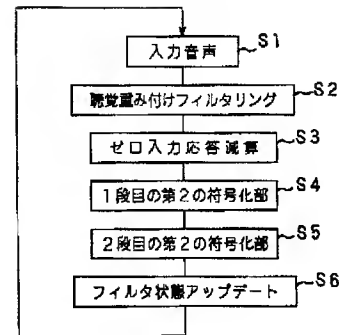
【図21】



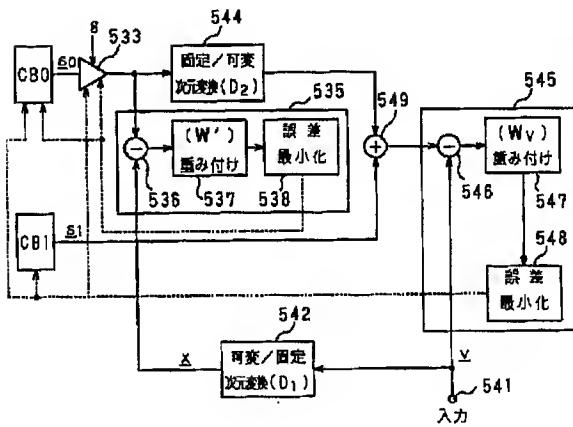
【図10】



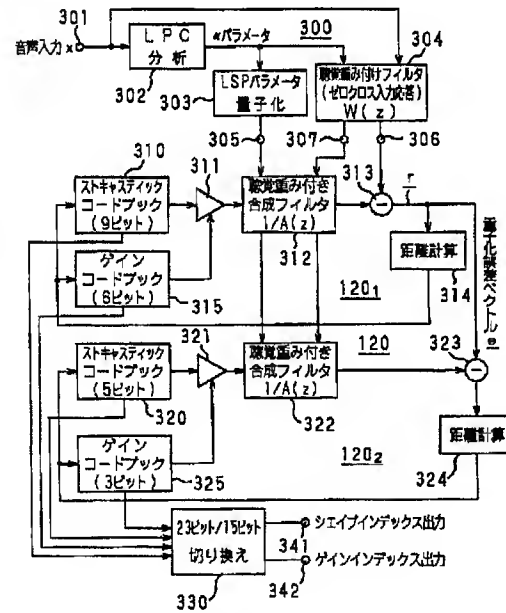
【図17】



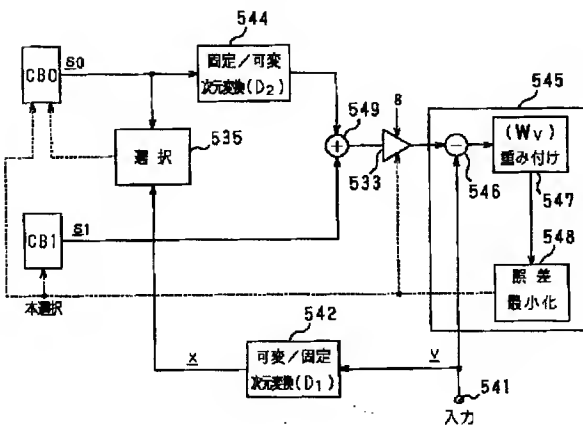
【図12】



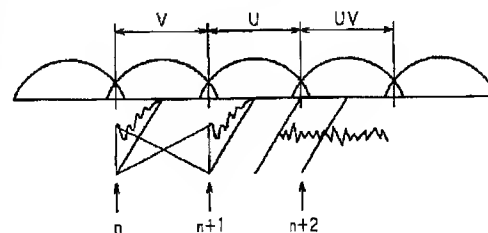
【図16】



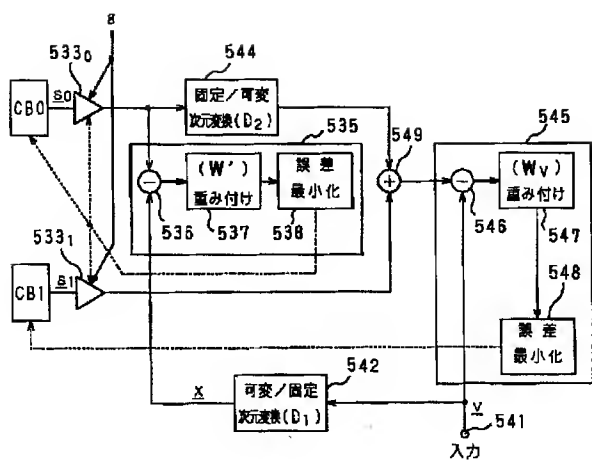
【図13】



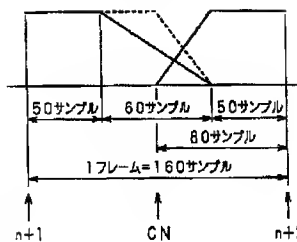
【図22】



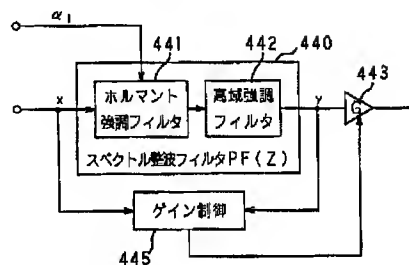
【図14】



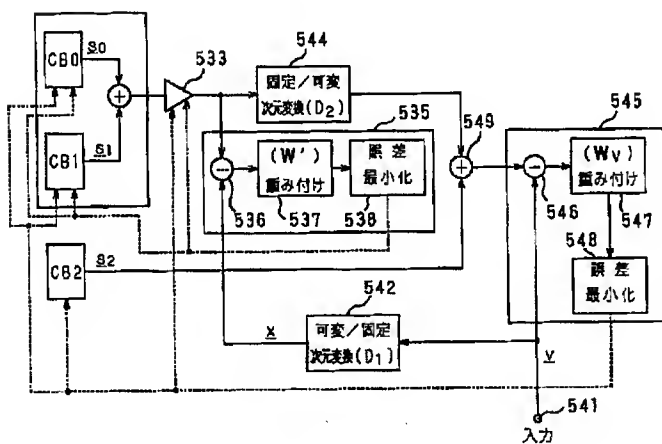
【図23】



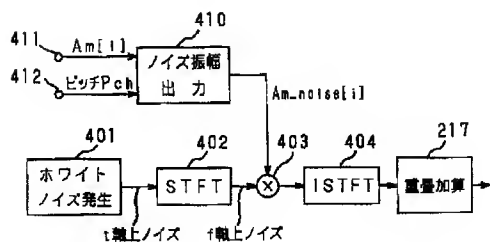
【図26】



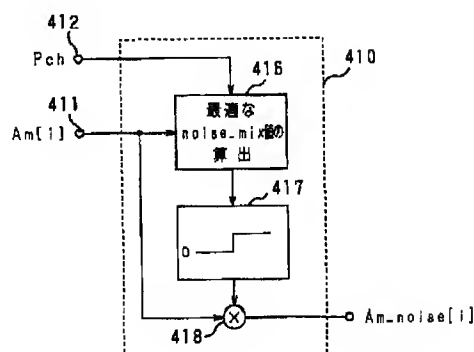
【図15】



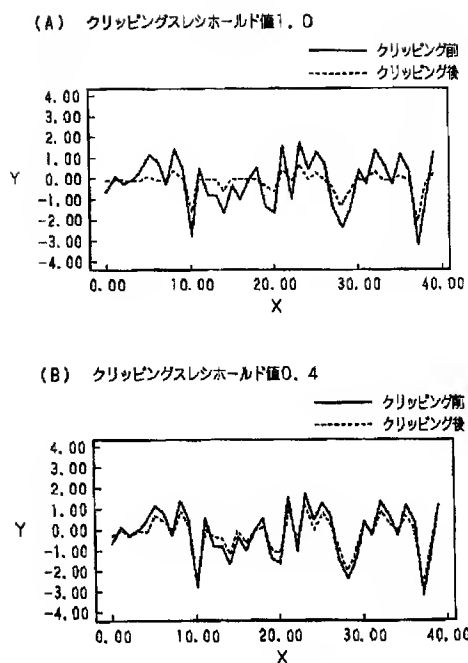
【図24】



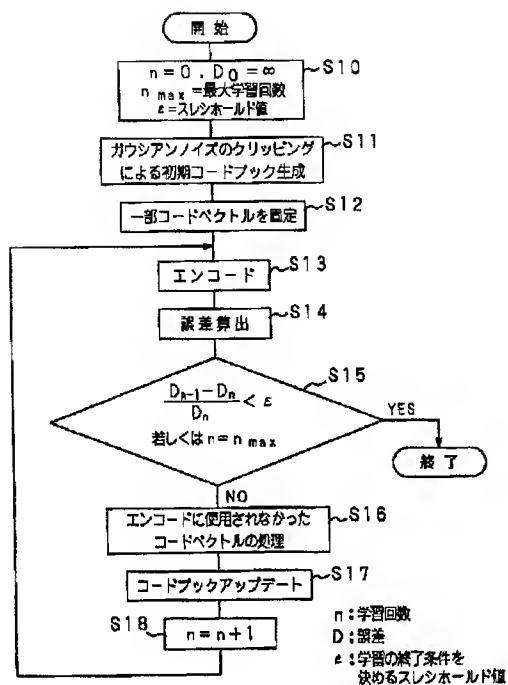
【図25】



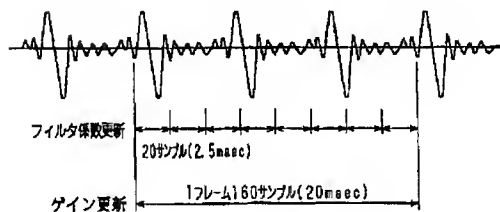
【図18】



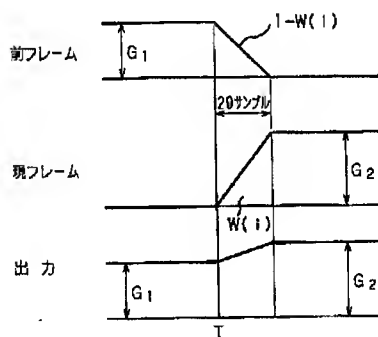
【図19】



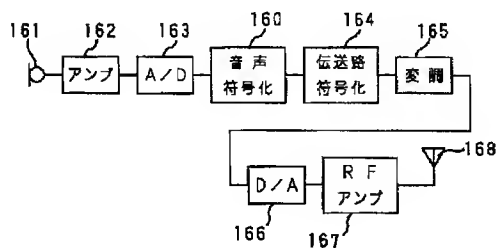
【図27】



【図28】



【図29】



【図30】

